59

One of skill in the art will appreciate that the order of the individual antigens within the fusion protein may be changed and that comparable activity would be expected provided each of the epitopes is still functionally available. In addition, truncated forms of the proteins containing active epitopes may be used in the construction of fusion proteins.

5

From the foregoing, it will be appreciated that, although specific embodiments of the invention have been described herein for the purpose of illustration, various modifications may be made without deviating from the spirit and scope of the invention.

60

SEQUENCE LISTING

- (1) GENERAL INFORMATION:
 - (i) APPLICANTS: Rued, Staven G. Skeiky, Yasir A.W. Dillen, Davin C. Campos-Neto, Antonio Noeghton, Raymond Vedvick, Thomas S. Twardzik, Deniel R. Lodes, Michael J.
 - (ii) TITLE OF INVENTION: COMPOUNDS AND METROES FOR IMMUNOTBERARY AND DIAGNOSIS OF TUBERCULODIS
 - (iii) NOMBER OF SECUENCES: 214
 - (iv) CORRESPONDENCE ACCRESS:
 - (A) ADDRESSEE: SEED and SERBY LLP
 - (B) STREET: 6300 Columbia Center, 701 First Avenue
 - (C) CITY: Seattle
 - (O) STAPE: Washington
 - (E) COUNTRY: USA (F) 31P: 98204-7092
 - (V) COMPUTER READABLE FORM:
 - (A) NEDIUM TYPE: Floory disk
 - B) COMPUTER: IBM PC commetible (C) OPERATING SYSTEM: PC-DOS/NS-DOS
 - (0) SOFTWARE: Patentin Release \$1.0, Version \$1.36
 - (will CUMBERT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE: 01-OCT-1997
 - (C) CLASSIFICATION:
 - (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Maki, David J. (B) REGISTRATION NUMBER: 31,392
 - (C) REFERENCE/DOCKET NUMBER: 210421,41107
 - (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (206) 622-4950
 - (B) TELEFAX: (206) 682-6631
- (2) INFORMATION FOR SEC ID NO:1:
 - (1) SECUENCE CHARACTERISTICS:
 - (A) LENGTS; 766 base pairs
 - (B) TYES: mucleig anid
 - (C) STRANDEDWESS: single (D) TG9OLOGY: linear
 - (xi) SECURNOU DESCRIPTION; SEC UP NO: 1:

61

| CGASGCACOS | GTAGTTTOMA | CCAAACGCRC | AATCONCOG | CAAACGAACG | GAAGAACACA | 6 |
|------------------|---------------|-----------------|-------------|------------|------------|-----|
| ACCATGAAGA | TESTSAAATO | SATCGCCGCA | GGTCTGACCG | coscoccesc | AATOGGOGCO | 3.2 |
| 9C70090C00 | GTGTGAGTTC | GATCATGGCT | 06000000366 | TOGTATACOA | GATGCAGCCO | 1.0 |
| Trespertee | GOGGGGGACT | GCCVTTGGAC | CCOCCATCCG | COCCTGACGT | CCCGACGGCC | 24 |
| OCCUACTICA | CCACCCTGCT | CARCAGOCTC | GOOGRECCCA | ACGIGICGTT | TGCGAACAAG | 30 |
| GGCAGTCTGG | TOGAGGGCGG | CATOGGGGGG | ACCGAGGGGG | GCATOGCGA | CCACAAGCTG | 36 |
| AAGRAGGCCG | CCGAGCACGG | GGATCTGCCG | CTGTCGTTCA | GCGTGACGAA | CATCCAGOGG | 42 |
| secceccecs | STTCGBOCAC | CGCCGACGTT | TOCCUTCTOGG | STOCGAAGOT | CTCGTCGCCO | 48 |
| GTCACGCAGA | ACGTCACGTT | COTTOAATCAA | GGCGGCTGUA | TGGTGTCACS | CGCATCGGCS | 54 |
| ATGGAGTTGC | TGEACGCCCC | AGGGNAACTO | ATTGGCGGGC | CHRITTCAGG | COSCISITOR | 508 |
| GCTACGCCGC | cceccasers | ACCCCTCCAT | GTCGARCACT | CGCGCGTGTA | GCACGGTGCG | 56 |
| OTHTOCGCAG | GGNCGCACGC | ACCROCCOST | GCANGCOGTC | CTCGASATAG | STOSTGNOTO | 72 |
| GNCACCAGNG | ANCACCCOOR | MNTCGMCMNT | TOTOGNIGHT | GNATGA | | 76 |
| AND THEFTS WAS A | EST ONLY ONLY | NY TEN SON LIVE | | | | |

(2) INFORMATION BOR SEQ ID NO. 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LEMOTH: 752 base paixs (B) TYPE: necleic acid

(C) STRANDEDWEDS: single (B) TOPOLOGY: linear

1xi) SEQUENCE DESCRIPTION: SEQ ID 80:3;

ATGUATUACO ATGACONTOS CUATGRACTO ACGGUNOSES CRACOTOCOS CITÓCOGOSCA 66 SACTICUICA CUGAGOTOGA CUTTUUTGUS CAAGOGGGTA CHGAGAGGCGC GGTOTUUTGGG 120 GIGGAAGGEC TODOCCOOG CICOGCGITG CIGGIAGICA AAGGAGGCCC CAACGCCGG 130 TOPOSETTOO TACTOGROSA ACCORTOROS TORCOTORES GECATOCOCA CARROACATA 240 ITTOTOGACG ACETSACOGT GAGCOGTCOC CATGOTGAAT TOUGGTTEGA AAACAACGAA 300 TYCAATGTCG TUGATGTCGG GAGTCTCARC GUCACCTACC TUGACCGCGA GEOCGTONAY 350 TOGOCCOTOR TOGOCARCON CORCORDOTO CACATOCOCA ACCTUCTOTI SOTOTICTIC 420 ACCGCACCCA AGCAAGGUGA GEATGACGGG ACTRODOGGG GCCCGTGAGC CCACCCGATA 483 GCCCCCCCCC GGCCGGGATG TOGATCGGGG CGGTCCTCCG ACCTGCTACG ACCGGATETY 940 COCTGATGIC CACCATORCC AAGAITOGAT TOTISGGAGG CITGAGGGTC NOONTGACCC 600 CCCCCCOGOCO CTCATTCHGO ONTHICCOCCH COTTICACCC CHIACCHACT CCCRCCCCCC 660

62

| TYGORAATIC NTICTYCHCI GCOONNAAAG GGACCHYFAN CITGCOGCTN GARANGGINA | 726 |
|---|---------|
| TOUNGEGOOD STEETINGAAN COCCHTOCOT OF | 750 |
| (2) INFORMATION FOR SEQ ID NO:3: (3) SEGUENCE CHARACTERICS: (4) LEMENT: 8:3 berg pairs (5) TYPE: nuclei ped (6) STRANEDURDE: single (D) TOPOLOGY: linear | |
| (xi) SEQUENCE DESCRIPTION: SEQ ID NO.3: | |
| CATATGCATC ACCATOAGGA TCACACTTCT AACCGCCCAG CGCGTCGGGGG GCGTCGAGGA | 60 |
| CCACGCGACA CCGGGCCCGA TCCATCTCCT AGCTTGAGTC TGGTCAGGCA TCGTCGTCAG | 120 |
| CAGCOCCATS CONTATOTTY STORTGERCY CASATATOGC GGCAATCCAA TETTECOSCCY | 180 |
| GCGGCCGGCG GTUCTGCAAA CTACTCCCCG AGGAATTTCG ACGTGCGCAT CAAGATCTTC | 240 |
| ATGCTGGTCA DESCRETCHT TITECTCTGT TOTTCGGGTG TGEMCMACGGC CGCGCCCAAG | 300 |
| ACCIACIGCO AGGAGTIGAA ASSCACCOAT ACCGGCCAGO CGIGCUAGAT INAAMIGICO | 366 |
| GACCOGGCCT ACRACATORA CATCAGOUTG COCAGITACIT ACCOCGAGOA GARGTOGGTG | 420 |
| SAAAATTACA TOUCOCAGAC SUGUGACAAG TTOUTCAGOI ODGUCAMATU STOMACHUCA | 4 18 (3 |
| CECGAAGCCC CCTACGAATT GAATATCACT TUBECCACAT ACCAGTCCGC GATACCGCCG | 540 |
| COTOGRACOC AGGOCOTOCT OCTCAMOGIC TACTREAACT CCOGCOGCAC GCACCCMACC | 600 |
| ACCACGTACA AGGCUTTOGA TEGGGGCCAG COUTATOGCA AGGCGAATCRO CTATGACACG | 660 |
| CTGTGGGAGG CTGACACCGA TCCGCTGCCA GTCCTCTTCC CCBTTGTTGC AACGTGAACT | 720 |
| GASCARCGCA GACCAGGACA ACMOGTATOG ATRICCOXXVM ARTGCCCCCCT TGGARCCCMG | 786 |
| TGAAATTATC ACAACTTCGC AGTCACHAAA MAA | 815 |
| (2) INFORMATION FOR SEQ 10 NO:4: | |
| (1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 447 base pairs (B) TYPE: nonleic acid (C) STRANGELMESS: single (D) TOPCLOGY: linear | |
| | |

(xi) SEQUENCE PERCRIPTION: SEQ ID NO:4:

CONTATGAAC ACOGCOGGGT CUCATAMOTT CONSCROTOC CAGGGTOGGC ACOGATICSC

| CAPTCCGATC | REGEROGESIA | TOXIGATUGE | SGGCCAGATC | CGATORAGIG | GGGGGTCACC | 120 |
|------------|---|--------------|-------------|--------------|------------|-----|
| CACCGFTCAT | ATCGGGCCTA | CCCCCTTCCT | occurrect | GTTOFCGACA | ACAACGGCAA | 186 |
| OGOOGGAGGA | GTCCAMCGCG | TOTOTOTA | CGCTCCGGCG | SCAMSTOTES | GCATCTCCAC | 240 |
| csecuncers | ATCACCGCGG | TOBACGGCGC | TOOGATOMAC | †COGCCACCG | CGATGGGGGA | 300 |
| COCCCTTARC | GGGCATCATC | CUGGTGACGT | CATCTCGGTG | AACTGGCAAA | CCAACTCGGG | 360 |
| CGGCACGCGT | ACAGGGAACG | TGACATTESC | CGAGGGACCU | COGGOOTGAT | TTOSTOSTGG | 420 |
| ATACCACCCG | cogecogeocc | AATTGGA | | | | 447 |
| (S) INFORM | ATION FUR S | EQ 10 80:5: | | | | |
| | (B) TYEB: D (C) STRABDE (D) TOPCLOS | CMESS: sing | le | | | |
| (x.i.) SI | EQUENCE DES | URIPTION: SI | PQ TD 90:5: | | | |
| GYCCCACTGC | ocroccesso | TATOTOGCCC | AGCAAATGTC | TOGENSOCIAC | CCAACGGAAT | 60 |
| COGGTCATOC | GACGTOGCAG | STTSTCGAMC | cosceaceae | agaactatog | GTCCATGCCT | 120 |
| AGCCCGGGGA | COGCGAGGGG | CORAATOGCS | CCACTGAGGA | GGCGGGGGAAT | TYGGCGGGGC | 180 |
| CCGGCGACGG | NGAGCGCCGG | AATGGCGCGA | GYSAGGAGGT | GSNCASTCAT | GUCCAGNGTG | 240 |
| ATCCAATCAA | CCTORATTCG | GNOTGROGGE | CCATTTGACA | ATCGAGGTAG | TGAGCGCAAA | 300 |
| TGAATGATGG | AAAACGGGBG | GMGAOGTCCC | BTGTTCTGGT | GGYGRTAGGT | GNOTONOTOG | 360 |
| ROTHGNOONT | ATCAGGATGT | TOTTOGNOSA | AARCTGATGE | CGAGGAACAG | GETETNICAX | 425 |
| NNANNCCHAR | GGNGTECNAN | CCCMMMMTCC | TOGNOGRANAT | CANANAGROG | nttgatgmga | 480 |
| NAAAAGGGTG | GANCAGNNNN | AANTNGNGGN | CONAANAANO | DIMMARKSHING | MAGNINGNI | 540 |
| MENTETTENC | ANDMINNETS | NNGNNGNNCN | NUNCAMENS | STNNBUGNAA | NNGGNTTN77 | 500 |
| MAAT | | | | | | 604 |

- (2) INFORMATION FOR SEQ TO ROIS:
 - (1) SEQUENCE CHARACTERISTICS:
 (R) LENGIR: 833 base pairs
 (B) TYPE: Gucleic sett
 (C) STRANDEDMESS: Single
 (D) TOPOLOGY: linear.

64

| 126 1 1 | SECTIONS | DESCRIPTIONS: | 2530 | 3.93 | N20 + 6 |
|---------|----------|---------------|------|------|---------|

| FFECANCTCG | AACCACCTCA | CTAAAGGGAA | CAAAAGCING | AGGTCCACCG | CGGTGGCGGC | 6 |
|------------|------------|------------|------------|------------|------------|-----|
| CGCTCTAGAA | CTAGTGEATH | YYYUKGGCTG | CACSAATYCG | SYNCOAGCAT | 7AGGACAGTC | 121 |
| TAACGGTGCT | GPTACGGTGA | TOSAATGACO | GACGACATCC | POCTGATCOA | CACCGACGAA | 100 |
| CGGGTGCGAA | COCTOACCOT | CAROCGGCUS | CAGTOCOGYA | ACCCCCTCTC | GGCGGCGCTA | 240 |
| CGGGATOGGT | TTTTCGCGGY | OTIGGYCGAC | GCCGAGGYCG | ACOACGACAT | CGACGTCGTC | 300 |
| ATCCTCACCG | GYGCGGATCC | GGTGTTCTGC | GCCGGACTGG | ACCICAACCI | ASCTGGCCGG | 360 |
| GCAGACÓGCG | CYGCOGGACA | TOTOACOGOO | Greececc | ATGACCAAGO | COGTGATOGG | 420 |
| COCGATUAAU | GUCGCCGCGG | TCACCGCCGG | GOTUGAACTG | SCSCTSTACT | SCGACATOCT | 480 |
| GATOGCOTCO | COCCACACA | GCTTCONCCA | CACCCAGGCU | CSSGTGSGSC | TOOTGCCCAC | 540 |
| crossgacte | AGTGTGTGCT | TGCCGCAAAA | COTOMICATO | GGNCTGGGCC | GOTGGATGAG | 600 |
| COTGACCGGC | GACTACCTGT | CCCTGACCGA | COC | | | 633 |
| | | | | | | |

(2) INFORMATION FOR SEQ ID NO!7:

US SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1362 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDMESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

OSSCACCACO CASCASEXAS TOSANTONTS AGAITTSTCA ACCATATEGA SCOCCECCO CUCOGCUGAG COGGCGCCC GGTCCCCGAG GTCTATGCUG AGGCCCGCCC CGAGTTCGGC 180 COGCIGOCOS ASCOSCIOSO CATOCITTOS COCCACGASO SACISSTICAS COCCOSCITO 240 CARACTITIC GOGAGACACY GOTGOTOGOC CARACTOLICSU GISGOCOCAA CGAAGACGTC 300 GCCGOCGCCG TOROGGCCAG CCTGCGCTGC CCCTGCTGCG TCGACGCACA (ACCACCATG 380 CTOTACODES CASSICARAS CONCACOSOS SOCIOCASTOS TESCOSOCAS ASCASOTEDO 420 SCCOSTGACO OGRACOCOCO GTATETOROS TORGOTOCAS GRACOGOCAC ACCOGOGOGA 480 CONCERNAC ECTTORICCE GRATETECCE ONDRANTACE TREGEROUSE SCHOMANTE San CACTTCATOG CAOGCCTSGI CCIGGIGCTG CIGGACGAAA CCIICCIGCC SGRICHCCCC 600 CSCOCCEAS AGCTERTGES OFSESCOST SCACTGGTGT TESCOCOSEAA GGTGCGCCC 660 GASCATOGGO COCGOCGOY: LACCOCGOGG CTOXASCOGO GAROSCISCO CERCGATOTO

CORDGRACGNO SGCCCCGGNG AGCCGGCCCC AACGCCGGATC GACCCGGGCCC TCCCCAGAGT

65

| SCATGGGGGAA | CACCGFCCGA | GOCCATAGCA | ACCOCCTTCC | DUSCRETORS | CCACCACCTO | 780 |
|-------------|--------------|-------------|-------------|------------|------------|------|
| SACACOGCGC | CGCACCTGCC | GCCACCGACT | CGTCAGGTGC | TOAGCORGET | CGTGGGGTCS | 840 |
| TUGCACGGCG | AGCCAATGOC | GATGAGCAGT | COCTGGAUGA | ANNANADAC | COCCOACCYC | 900 |
| EDUGUUSACC | TGCACGCGCC | CACCCCGTCTT | GCCCTGCTGA | CCSCCCTSGC | OCCGCATCAG | 360 |
| GTGACCGACG | ACQACGTOSC | CGCGGCCCCGA | Tecengeros | ACACCGATGG | GGCGCTGGTT | 1020 |
| SACGCCCTTGG | CCARRECCEC | CTYCACCGCC | @CGCGCGCA | TOSGCACCTS | GATCGGCGCC | 1085 |
| GOCGCCGAGG | SCHROGRATC | GCGGCAAAAC | CCCACTGCAT | GAGTGTGCGC | SCOCTSTCOR | 1140 |
| TAGGGTGTCA | TOSCIEGOOC | GAGGGATCTC | GCGGCGGCCGA | ACGGAGGTGG | CGACACAGGT | 1200 |
| GGAAGCTGCG | CCCACTGGCT | TGCGCCCCAA | COCCOTCOTS | SGCGTTUGGT | TOGGCGCACT | 1260 |
| GGCCGATCAG | GPCGGCGCCG | SCCCTTGGCC | GAAGGTCCAG | CTCAACGTGC | DUTCACCGAA | 1320 |
| GGACCOGACG | GTCACCGGGG | GtdAccerse | GCHCCCAAGG | AA | | 1362 |
| (2) INFORM | UTTON FOR SE | WE TO MOTE | | | | |

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1458 bese pairs

(B) TYPE: monleic acid (C) STRANDEDNESS: single

(D) TOPOLOGY: Linear

(x1) SEQUENCE DESCRIPTION: SEQ 10 NO:4:

GUGACGACCO DGATATGCOG GOCACCGTAG CGAAAGCCGT CGCCGACGCA CTCGGGGGGGC 60 GTATORCTCC CGTTGAGGAC ATTCARGACT GCGTGGAGGC CCGRCTGGGG GAACLYYDTWC 126 TOGATGACGI GUUUGTGIT TACATCATCI AKXOGCACO GOGCGCGAG CIGCGCGACOA 186 CTAAGGCCTT GCTCGCCGTT CGCGACGACT TAAAGCTGAG CTTGGCGGCC GTGACGGTAC 240 TORGOJAGOS CTATOTOCTO CACGACGAGO AGGGCCOGGO GGACGAGTUG ACOGGGGAGO 305 TGATGGACCG ATCGGCGCC TGTGTCGCGC CGGCCGAGGA CCAGTATGAG CCGGGCCCCCT 360 CGAGGOGGFG GGOCHASCGG TTOGCCACGC TATTACGCAA CCFGGAATTC GTGCCGAATT OSCOCRACTT GATGARCTUT GGCACCGACC TSOGACTSCT COCCGCCTGT TTTCTTCTCC 480 CSATTGAGGA TYCOCTSCAR TCGATCTTTG CGAGGCTGGG ACAGGCGRCC GAGGTGCAGG 540 GENCTHORAGE USKNIACUEGA TAIGUSTTUA DOCAUCTUGE ACCOCCUEGE GATOGGETES 600 CCTCCACGGG CGGCACGGCC AGCSGACCGR ESPCSTTTCT ACGGCTGTAT GACAGTGCCC CGGGTGTGGT CTCCATGGGC GGTCGCCGGC GTGGCGCCTG TAPGACTGTG CTTGATGTGT OGCACCOCCA TATCTETGAT TTUCTCACCG SCAAGGCCGA ATCCCCCAGC GAGCTCCCCC 780

PCT/US97/18293

| ATTTCAACCT | ATCCUTTGGT | GTGACUGACG | CETTOCTGGG | OGCCGTCGAA | COCAACGGCC | 840 |
|------------|------------|-------------|-------------|------------|------------|------|
| TACACOGOCT | GGTCAATCCG | CXIAACCGGCA | AGATCGTCGC | OCCURTOCOC | GCCGCCGAGC | 900 |
| TOTTCGACGC | CATCTGCAAA | GCCGCGCACU | COSGTGSCGA | recongents | GIGITTOTOS | 960 |
| ACACGATCAA | TAGGOCAAAC | occerracea | GGAGAGGGCCG | CATOGAGGOG | ACCAACCCGT | 1026 |
| GCGGGGAGGT | OCCACTOCTS | CCTTACGAGT | CATGTARTOT | CGGCTCGATC | AACCTCGCCC | 1680 |
| GGATGCTCGC | CGACGGTCGC | OTOGROTOKIG | ACCOGUTOGA | GGAGGTCGCC | SCHOOLS | 1149 |
| RECOGNICCT | TOATGACGIC | ATCGATGTCA | GCCCCTACCC | CTTCCCCGAA | etgegtgaeg | 3206 |
| CGGCCCGCGC | CACCOGCAAG | ATOSGSCTOS | GAGTCATGGG | TTTGGCGGAA | CTGCTTGCCG | 1260 |
| CACIGGGTAT | TOCGTACGAC | agtgaagaag | CONTROSOTT | AGCCACCOGG | CTCATGCGTC | 1320 |
| GCATACAGCA | SGCSSCGCAC | ACCCCATORC | GGAGGCFGGC | CGAAGACCOC | GOCGCATTCC | 1386 |
| CGGCKTTCAC | CGATAGCCGG | TTCGCGCGCT | CGGGCCCGAG | SCGCAACGCA | CREGVCRCCT | 1440 |
| COGTEGETCE | GACGGGCA | | | | | 1458 |
| | | | | | | |

(2) INFORMATION FOR SEQ 10 MO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LEBCTH: 862 base pairs (B) TYPE: nucleic acid

(R) TYPE: nucleic acid (C) STRANDEONESS: single

(D) TOPOLOGY: Linear

(Mi) SEQUENCE DESCRIPTION: SEQ 10 NO:9:

ACCOTOTAAT COTOCTCCAT CTOCKACCOU STCCCCCCT ACCTACCUAC ATCTWCTCCC 60 9909CA9G96 GUTGGOCCT9 99CATCGC56 TCGTCGTA6T C69GATCSCG GYGGCCATCG TURTUGUCTT USTOGACAGO ROUGUUGGFS COAAACOGGT CAGUGUUGAC AAGOCGGTUT 186 OCCCCAGAG CENTECEGOS TERCORREAS CECHAGOACO CEAGCOGACE AGOCAARCER 240 AASSTAACGO CGCCCCGGCC CCCCCGCTAGG GCCAAAACCC CGADACACCC ACGCCCACCG 300 COSCECTORA GOOGLOSKING STRUTCHAGO AAGGEGACGA TEGOXICGAT TOGACCUTGG 360 CCGTCAAAGG TTTGACCAAC GCRICKEIAGT ACTACGTCGG CCACCAGCCG AAGTTCACCA 420 TOOTOGPCAC CAACATCOGC CYGGTGTCCT GTAAACGCGA CUTTGGGGGCC GCGGTGTTGG 489 COGCCTACGT TRACTOROUG GACAACAAGC GETTGTEGTC CAACUTGGAC TGCSCGCCCT 540 CHARTHAGAC OCTGOTCARG ACGTTTTCCC CCGGTGABCA GGTARCOACC GCGGTGALCT 500 SUBACCOGGAT GOGATOGGGG COGCGCTGCC CATTGCCGCG GCCGGCGATC GGCCCCGCCA 660 OCTACATOT CETEFIACAA CFGGGCAAYC TEGGCTCGCT GCCGGTTGCG TTCATGCTGA

| 97 | |
|---|-----|
| ATCASCESCE SCENCESCE SSSCESSTAC CONCILCOGO TOCASCECAS SOSCOPOGO | 780 |
| CCCASTCTOC OGCOCARGOS CGATAATTAT TGATCOCTCA TEGTCOATTC CCCCAGCTGT | 840 |
| GACRACCCCT CGCCTC97CC CG | 862 |
| (2) INFORMATION FOR SEQ IS MO:10: | |
| (i) SEQUENCE CRARACTERISTICS: (A) LENGTH: 622 base pairs (B) TYPE: nucleic scid (U) STRANDENDESS: single (G) TOPOLOTY: Linear | |
| (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10: | |
| TTGATCAGCA GEGGKAAGGE STEACATGES TOSCIGGOTG TSCAGGTGAC CAATGACAAA | €0 |
| GACACCCCCC CCCCCAAGAT CETOGAAGIA STGGTCN/TE GTGCTGCCCC CAACCCTGGA | 120 |
| GTGCCGAAGG GOGTOTTGT CRCCAAGGTC GACGACOGCC CGRTCAACAC CGCGGACGCG | 180 |
| TTGGTTGCCG COSTGCGGTC CANACOGCCG SECGCCACUS TYSCCGCBASC CTTTCACGAT | 240 |
| SUTTOGRADO GTAGUCOGAS AGTOCAAGTO ACUCTOSGGA AGGOGGASCA GTOATGAAGG | 308 |
| TESCENCICA CYGYYCAAAG CYGGGATATA COSTGGCACC CATOGAACAG CGTSCSGAGY | 360 |
| TOUTGOTTOG DOGGGCACTT GIVGTCGTCG TYGACGATCG CNCGGCGCAC GGCGATGAAG | 420 |
| ACCACAGCOG GCCGCTTGTC ACCGAGCTGC TCACCGAGGC CSGCTTTGTT GTCGACGGCG | 400 |
| TESTESCOST STEESCOURC GAUSTESAGA TOCKARATOS SCHEARCACH SOCKTOATOS | 540 |
| GESGGGTGGR CUTGUTGGTG TEGGTCUGCG GGACCGGRGT GACGUCTOCC CATGTCACCC | 600 |
| CGGAAGECAC COONGACATT CT | 622 |
| (2) INFORMATION FOR SEG ID BO:11: (1) SEQUENCE CRARACTERISTICS: (1) LEMSTH; 1200 Dame pairs (2) TYPE: nacleic scid (CI STRANDERSESS: single (D) FORDLOGY: linear | |
| (NI) CEQUENCE DESCRIPTION: SEQ 10 NO.11: GUCGURGEST TRANSCRIPT GUCGECOSTS ACMITSTOF TRANSCRIP COCCOSTOC | 66 |
| ACCAACASCT COTOGTCASG CSCAGSCSGA ACGTCTGGGT COGTGCACTG CGGCGGCAAG | 120 |
| ANGGROCTEU ACTURNOUS CTOGROUNG CARGANANTO CONTOGNACA GTTOSTETAT | 180 |
| | |

68

| SCETACOTOC | GATOGTGOOD | GUGUTACACC | TTGGACTAGA | ACGCCAACGC | GTCCGGTGCC | 24 |
|------------|------------|------------|------------|-------------|------------|------|
| GUGGTGACCC | AGTTTGTCAA | CAACGAAACC | GATTTOGCCG | GCTCGGATGT | CCCCTTGAAT | 30 |
| COSTORACCO | STCAACCTGA | COGGTGGGG | GAGCGGTGGG | GTTCCCCGGC | ATGGGACCTG | 36 |
| CCGACOGTGT | TOSGCOCCAT | OGOGATOACC | TACARTATCA | AGGGCGTGAG | CACGCTGAAT | 429 |
| CTTGACGGAC | CCACTACCGC | CAAGATTTTC | AACGGCACCA | TOACCGTGTG | GAATGATCCA | £81 |
| CAGATCCAAG | CCCTCRACTC | CGGCACCGAC | CTGCCGCGAA | CACCIGATING | COTTATCTTC | 544 |
| CSCAGCGACA | AGTOOGGTAC | GTCGGACAAC | TTCCAGAAAT | ACCTOGACGG | TGTATOCRAC | 501 |
| 990GCGTGGG | GCAAAGGCGC | CAGCGAAACG | TTCAGGGGGG | cogresses | COGCGCCAGC | 560 |
| GGGAACAACG | GAACGTCGGC | CCTACTGCAG | ACGACCGACG | GGTCGATCAC | CTACAACGAG | 720 |
| TOGTCGTTTG | CGGTGGGTAA | GCAGTTGAAC | ATGGCCCAGA | TCATCACGTC | GGCGGGTCCC | 780 |
| SATECACTOS | CGATCACCRC | CONCTORNIC | OCTANGACAA | PGGGGGGGGG | CAAGAYCAYG | 841 |
| GGACAAGGCA | ACGACCTGOT | ATTGGACACG | regregreer | ACAGAGCCAC | CCASCCTGGC | 900 |
| TOTTACCOGA | TOGTGOTOGO | GACCTATGAG | ATOGTOTOCT | CGAAATACOO | GGATGCGACG | 960 |
| ACCESTACTS | COSTANGGGC | CTTTATGCAA | OCCGCGATTG | GTCCAGGCCA | AGAAGGCCTG | 1926 |
| GACCAATACE | GCTCCATTCC | GTTGCCCAAA | TORTTOCAAG | CARAATTOOC | occccceere | 1088 |
| AATGCTATTY | CTTGACCTAS | TGAAGGGAAT | TOGACGGTGA | GCGATGCCGT | TCCCCAGGTA | 3140 |
| SECTOCAAT | TTGGGCCGTA | TCASCTATTS | COCCTECTOS | GCCGAGGCCG | GATGGGCGAG | 1200 |
| | | | | | | |

(2) INFORMATION FOR SEQ ID NO:12:

ill SEQUERCE CHARACTERISTICS:

(A) LENGTH: 1155 base pairs (B) TYPE: numbers acid

(C) GTRANDEDNESS: single (D) TOPOLOGY: linear

(xi) ABQUENCE DESCRIPTION: SEG 19 NO:12:

GCARGCASCI SCAGSTOWYS CITOTYCGACG ARCIGGGCRI GCCGAAGACC AAACGCACCA 50 AGACCOCCTA CACCACRGAT SCCCACCCCC TGCAGTCGTT GTTCCGRCAAG ACCGGGCATC \$20 COTTECTORA ACATOTOCTO GOODACDOOG ACGTOADOOG GUTCAAGGFO ACCSTEGACU 180 SGTTSCTCCA AGGSTTSGCC SCCGACGGCC SCATGUACAC CACGTTCARC CAGRCGATCS 240 CUGUSACUGG CUSSUTUTUU TUSACUSAAU CUAACUTGUA GAACATUCUU HTYUGAAGUS 300 ACCCCGCCCC GOOGATCCCC GACCCCTTCC TOSTCCCCCGA CCCTTACCCC GACTCATGA 360 COGCOGACTA CAGCOGGATE GAGATGCGGA TOATGGGGCA COTGTCCGGG CACCAGGGGC 120

| TCATCHAGGC | GTTCAACACC | 909GAGGACC | TOTATTUGET | COTORCIPIC | COGGTGTTCG | 483 |
|------------|-------------|-------------|-------------|------------|-------------|------|
| STGTGECCAT | COACGAGGTC | ACCORCGNOT | 10000000000 | GGTCAAGGCG | ATGTECTACQ | 540 |
| GGCTGGTTTA | OGGGTTGAGC | OCCTACGGGC | TOTOCOCAGOA | GTTGRAAATC | TCCACCGAGG | 600 |
| AAGCUAACGA | CONCATOGAC | CONTATTON | COCCATTOGG | caggaraced | ONCINCETEC | 680 |
| GOGCCGTAG7 | CGAGCGGGCC | CCCCAGGACG | SCYACACCTC | GACGGTGCTG | ascogresce | 720 |
| 6CTACCTGOC | OGAGCTUGAC | AGCAGCAACC | GTCAAGTGCG | GGAGGCCGCC | GAGC GGGCCG | 780 |
| COCTEAACCC | GCCGATCCAG | GCCAGCCCGG | COGACATOAT | CAAGGTGGCC | ATGATOCAGG | 840 |
| TEGREARGGE | OCTOANCOAG | ocacagengs | CGTCGCGCAT | 6CFCCTGChG | CTCCACGACG | 900 |
| AGCTSCTOTT | CGAAATCCCC | OCCOGCTGAAC | GCGAGCGGGT | COAGGCCCTG | GTGCGCGACA | 960 |
| AGATGGGCGG | CGCTTACCON | CTCGACGTCC | OGCTGGAGGT | STOSSTSON | TACGGCCGCA | 1020 |
| GCTGGGACGC | OGOGRAGACAC | TONOTOCOOR | COSTGCATOR | GGGGGGGAA | TTOGGCGATT | 1080 |
| TYPOCCOCCT | GASTYCACGC | TOSCOCRAF | COGGACCGAG | TTTGTCCAGC | STOTACCEST | 1146 |
| OGAGTAGOOT | CGTCA | | | | | 1155 |
| | | | | | | |

PCT/US97/18293

120

180

420

(2) INFORMATION FOR SEC 10 NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1771 base pairs

(B) TYPE: nucleic acid

(C) STRANOEDMESS: single (D) YOPOLOGY: linear

TRIL SECURIOR DESCRIPTION: SEC ID NO:13:

GAGGGCCCTC TCCTCTTTGA ACGCTTTTAC CURTCGGCAT C/GCACGGCC GTTYGCCGGCT 360 TOGGGCCTCG GGTTGGCGAT CCTCARACAG GTGGTGCTCA ACCACGGGGG ATTGCTGCGG ATCGARGACA COGACCURGO OGGOCAGCOU COTGGARCET OGATTIACET COTGOTTON GOCCOTOGGA TOCCGATTOC GCAGCTTCCC GATGCGACGG CTGGCGCTCG GAGGGCGGAC 240 AFCGAGAACT CTCGGGGGTC GGGGAACGTT ATCTCAGTGG AATCTCAGTC CACGCGGGCA 300 ACCTAGITEST SCACITACTS TIBARASCOR CACCDATGOS ASTOCACCOR TOGOCAMONT 366 GEOCOGAGTA GTOSECCTAE TACAGGAAGA GCAACUTAGC GACATGACGA ATCACCCACG CTATTCCCCA COGCOGCAGE AGCCOGGAAC CCCASCTTAT GCTCAGGGGC AGCAGCAAAC 480 GTACAGOCAS CASTICGACT GGCGTTACCC ACCETOCOOG CCCCOSCAGC CRACCCAGTA COGTUARCOU TROGRAGOST ISSUIGGTAC COSCUEGOST CIGRIACOTS SOSTSATIVO SOR GACCATGAGS OCCCCTCCTG SGRIGGTTOG CCAACGCECT CCTGCACGCA TGTTGGCCAT 560

| COGCOCOGTO | ACGATAGC60 | TSGTGTCCGC | CGGCATCGGC | 600808666 | CATCOCTGGE | 720 |
|------------|-------------|-------------|-------------|------------------|-------------|-------|
| CEGETTCAAC | COGGCACCCS | CUGGOCCCAG | CGGCGGCCCA | GIGGETGECA | coscacaca | 786 |
| AAGCATCCCC | GCAGCAAACA | TGCCGCCGGG | GTCGGTCGAA | CAGGTGSCCS | CCAAGGTGGT | 840 |
| GCCCAGTGTC | STCATSTINGS | AAACCGATCT | OSCCISCOAG | TORGREGACE | CCTYXXXXXX | 900 |
| CATTCIGICT | GCCGAGGGGC | TONTOTTGAC | CARCAMOCAC | GEGATCOCGO | OBGCCGCCAA | 960 |
| GCCTCCCCTG | SGCAGTCCGC | CGCCGAAAAC | GACGGTAACC | TICTOTEACE | GCCCGACCGC | 1020 |
| ACCOPTUACE | GTGGTGGGGG | CTGACCCCAC | CASTGATATO | GCCGTCGTCC | GTGTTCAGGG | 1080 |
| COTCTCCGGG | CYCACOUCUA | revecesses | TTCCTCCTCG | GROOTGAGGG | TOGGTCAGCC | 1140 |
| GGTGCTGGGG | ATCGGGTCGC | COUNCUSTRY | SGAGGGCACC | OTGACCACGG | GGATCOTCAG | 1200 |
| COCTCTCAAC | CGTCCAGTGT | CGACGACOGO | CGAGGCCGGC | RACCAGRACA | COSTSCTOGA | 1260 |
| OGCCATTORS | NEOGREGOES | CONTONACCO | COGTANCTOC | GEGGGGGGGGC | TOOTGAACAT | 1320 |
| GAACGETCAA | CTCCTCGGAG | TOALCTOGGC | CATTGCCACG | Official actions | ACTOAGOOGA | 1.386 |
| TGCGCAGAGC | SECTOGATOR | GTCTCGGTTT | TGCGATTOCA | OTCGACCAGG | COMMICCICAT | 1440 |
| OGCCGACGAG | TTGATCAGCA | COSSICAMBEC | GTCACATGCC | TOCCTGGGTG | TOMAGGTGAC | 1500 |
| Cartgacaaa | GACACCCCGG | GCGCCAAGAT | CÉTUGAAGTA | STORKER | GTGCTGCCCC | 3569 |
| GAACGCTGGA | GTGCCSANGG | GCGTCGTTGT | CRCCAAGGTC | RACGACUGCC | CGATCABCAG | 1620 |
| COGGACGOG | TIGGTTGCCG | COGTSCOCTO | CAARGENGEEG | GGGGGCACCG | TGGCGCTAAC | 1660 |
| CTITCAGGAI | CCCTCGGGGG | GTAGCCGCAC | AGTECAAGTC | Accercance | AGGEGGAGGA | 1.740 |
| otgatgaagg | TYGGGGGCCA | GFGTTCAXAG | C | | | 1771 |
| | | | | | | |

12) INFORMATION FOR SEC ID NO:14:

(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1058 base pairs

(8) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOFOLOGY: linear

(xi) SECONNCE DESCRIPTION: SEC ID NG:14)

CYCLACCECC GTGGGGGGGG CTCTAGAACT AGTGGATCCC CCGGGCTGCA GGAATTCGGC 60 AUGASGATOR CARREDGESS STRUTCHARD COSCOSCOGO GRANDTATOS STOCKTOCCT 120 AGCCCGGCGA CGCCGAGCGC CGCAATCCCG CGAGTGAGGA GGCGGGCÄAT TYGGCGGCGC 190 COSSOGACGO CORSCIOCOS ARTOGUECGA STORGUAGGO GOGCAGICAI SECCAGOSTO 240 ATCCARTORA COTSCRITCS OCCTSCSSOC CONTITUACA ATCCASCTAG TRAGCROMA 300

| TGAATGATGG | AAAACGGGCG | GTGACGTCCG | CIGITCIGAL | GGTGCTM9GT | GUCTGCCTYSG | 360 |
|------------|-------------|-------------|-------------|------------|-------------|------|
| COTTOTOCCT | ATCAGGATGT | TOTTOGOCCUS | AACCTGATGC | CGACGAACAG | GGTGTTCCCG | 420 |
| TGAGCCCGAC | GGCGTCCGAC | COCOCOCTOC | YOGOGGAGAT | CAGGCAGTCS | CTTGATGCGA | 486 |
| CAAAAGGGTT | GACCAGCSITS | CACGTAGGGG | TOGGAACAAC | CESCAAAGTC | GACIAGCTTOC | 540 |
| TEGUTATTAC | CAGTGCCGAT | GTOGACGTCC | GGGCCAATCC | COTOGOGGCA | AAGGGCGTAT | 600 |
| GCACCTACAA | CGACGAGCAG | GGTGTCCCOT | TTCSGGTACA | AGGUSACAAC | ATCTCGGTQA | 988 |
| AACTGTTCGA | CSACTSGAGC | AATCTC30CF | CGATTTCTGA | ACTGTCAACT | TEACGCGTGC | 720 |
| TOGATOCTEC | CGCTGGGGTG | ACCORDITION | TOTOCOGREGA | CACGRAGOTO | CANDODDANAG | 780 |
| GTACCGAAGT | GATAGACGGA | ATTTCGACCA | CCAANATCAC | COGGACGATC | CCCGCGAGCT | 840 |
| CTCTCAAGAT | COTTGATCCT | GGCGCCAAGA | GTGCAAGGCC | GGCGACCETG | TOGATTGCCC | 900 |
| AGGACGGCTC | GCACCACCTC | GTUCGAGCGA | GCATCGROOT | COCATCCAGG | TEGATTCAGE | 960 |
| TUACGCAGTO | GARATOGAAC | GAACCOGTCA | ACCTORACTA | GGCCGAAGTT | GCSTCGACGC | 1020 |
| GTTGNTCGAA | ACRCCCTTGT | GAACGGTGTC | AACIGNAC | | | 1058 |
| | | | | | | |

(2) INFORMATION FOR SEC 15 MO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 542 hase pairs
 - (B) TYPE: nucleic dold (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

ixi) SEQUENCE DESCRIPTION: SEQ TO NO:15:

GAATTOOGCA CGAGAGGTGA TCGACATCAT GGGGNCCAGC CCCACATCCT GGGRACAGGC 88 SOCGGCGGAG GCGGTCCAGE GGGCGCGGGA TAGGCTCGAT CACATCGGCG TOGUTCKGGT 125 CATTGAGCAG GACATGGCCG TOGACAGCCC CCGCAAGATC ACCTACCCCA TCAAGCTCGA 180 AGTGTCGTTC AAGRTSAGGC COOCCCAACC GCGCTAGCAC GGGCCGGCGA GCAAGACGCA 2531 ARATOGOROG STETSOSSETT GATFOSTSOS ATTTESTSTO TOUTOGOOGA GRUCTACOAG 300 SCOCHOCCEA SOTEESCOTE CTSCOSTATE CAGGESTOCA TOOCGATTEE GEOGGESCO 360 CODGADTTAN TECTTOGOGT CGACCOGANC TEGENORICO SCOGGNERICO TERTTORICA 420 CONTRECCAS COCCITORES ECUSACITEC COGASGARAC STOCTECCAS OCCUPITAGGA 480 AKTUTCUSTA ANGANOSGIG CIGACOGACT CINCCTACGC CCICAGIGCA GCCAGOGAGC 546 542

(2) INFORMATION FOR SEQ ID MO:16:

| (4) | SEOD | SNOR | CHAI | RACTS | CRIST | 1000 |
|-----|------|------|------|-------|-------|-------|
| | | | | | | pairs |
| | | TYPE | | | | |
| | (0) | STRA | NOUS | WES! |): sl | ngle |
| | (D) | TOPO | 200 | 1: 1: | near | |

(xi) SEQUENCE DESCRIPTION: SEQ ID NO.16:

| cacraccacu | CUCKCLYCC6 | TIGCCCCCAT | TOCCGCCCTC | GCOGATCAGC | TRUSCATOGO | 60 |
|---------------|-------------|--------------|---------------|-------------|-------------|-----|
| CACCATCACC | GCCTTTGCCG | CCGGCACCAC | CSGTGGCGCC | GREECCOOLE | ATGCCACCGC | 120 |
| ETOACCCTUG | ccaccacacac | CHOCATTOCC | ATWOMSCACE | 0000000000 | CACCGTTACC | 185 |
| XCCGTCGCCA | ccarcaccae | cccreecert | TOAGGCCOGG | GAGGCCGAAT | GAACUECOSO | 240 |
| CAMBRECOCC | 6000604000 | TYGECGECTT | TTUCGGEEGE | 0000000000 | CCGCCAATTG | 300 |
| CGAAACAGOC | AMGCACCOTT | SCCSCCAGCC | 00900000007 | TAACGGCGCT | acceeecec | 369 |
| CADDOOMS | CCCCCATTAC | CGCCGTTCCC | CTTCSGTGCC | COGCCGTTAC | excescesee | 420 |
| PTTEGCCGCC | AATATTCGGC | 999GACCGCC | AGACOCICOU | GGGCCACCAT | TOCCUCAGG | 180 |
| DACCGAAACA | ACAGCCCAAC | garaceaeee | GOCCGGCGT | TEGGOOGGAT | CACCINICONT | 540 |
| CACCGCCAG | CACCOCCOTT | ARTOTTTATO | AACCCOGTAC | ESCUMOCOCO | GCCCCTATTS | 600 |
| 200808093 | SAGNGCG19C | COGCCGGCGC | COCCAACGCC | CAAAAGOCOG | GGGTTGCCAC | 666 |
| 2900CCCCCCC | SKACCCACCS | GTCCCGCCGA | TOOCCCCGTT | 9000000000 | COSCOGECAT | 120 |
| eggeggetoet | GAAGCUGTTA | GCGCCGGTTC | CATCAGGTTYCC) | GGCGGTGGCYS | CONTESCOSC | 766 |
| 16GC000000 | GTTGCCGTAC | AGCCACCCCC | CGGTGGCGCC | STIGOOGCCA | TTGCCGCCAT | 840 |
| racescest? | GCCGCCATTG | cogcognition | GGCGGCCACC | OCCUGANTING | ccaccacce | 900 |
| secretarions. | nan | | | | | 013 |

(2) INFORMATION FOR SEC IN NOTITE

(i) SEQUENCE CHRACTERISTICS:
(A) LENTH: 1872 base pairs
(B) TYPE: syciet; soid
(C) STWAMBENNESS: single
(D) TOPOLOGY: linear

(WE) SEQUENCE DESCRIPTION: SEQ ID NO:11:

| GACTACGTTG | gtgtagaaaa | AFCCTGCCGC | CCGGACCCTT | AAGECTEKKA | CARTETETGA | 68 |
|------------|------------|------------|------------|------------|------------|-----|
| TAGCTACOCO | GACACAGGAG | GTTAGGGGAT | GAGCAATTCG | COCCOCCCC | CACTUAGGTG | 130 |

(2) INFORMATION FOR SEC 10 BO:18:

WO 98/16646 PCT/0897/18293

73

STOATSGITS CTGAGGGTGC TGGCTGCGGT CNGGCTGGGC CTGGCCACGG CACCAGGTA 180 9008600006 COSSCCTTOT COCAGGACCO GTTCGCCGAC TECCCCGCCC TGCCCCTCGA 240 CCCCTCCCC ATOSTCCCCC AASTCCCCCC ACAGGGGCTC AACATCAACA CCAAACTGGG 300 CTACAACAAC GCCGTGGGCG CCGGGACCGG CATCGTCATC GAZCCCAACG GTGTCGTGCT 350 SACCAMERAC CACCINATES CORRECCED CONCATENAS CONTRARES TOCCONOCOS 400 COMMANDETAC CHECKEGATG TECHTECOTA TENECOCACE CAGGATOTEG CAGTOCTACA 689 SCIECCECCT SCIEGOSSCO TECCETOSEC SCIEGATOSET OCCESCOTOS CONTIGOTOR 540 GCCCGTCGTC GCGATGGCCA ACAGCGGTGG GCAGGGCGGGA ACGCCCCGTG CGGTGCCTGG 600 CASSOTSOTO SOSCITOSSOO MAACONTSCA SUCCIOSIAT TOSCITAROS SINGOSTAGA 660 SACAPTGAAC GOGTTGATOU AGTTCGATGC CGCAATCCAC CCCGGTGATT CGGGGGGGCC 720 COTTOTANC CHRISTAGGAC AGGINSTUDE TATORACACO GOOGOTTOOG ATAACTTOCA 786 GCTSTOCUAG GGTGCGCAGG GATTCGCUAT TOCGATCGGG CAGGGGGATGG CAATCGCGGG 820 CCARATCUCA TOGGGTOGOU GETCACCOAC CUTTONTATO GEGOCTACOS CUTTOCTORO 900 CITGOGTOTT GTCGACAACA ACGGCAACUG CGCACGGGTC CAACGCGTCG TCGGAAGCGC 250 TOCGGOGGGA ACTOTOGGCA TOTOCACCDS DGACGTGATO ACCGCOGTOG ACGGOGGTCC 1000 GATCARCTOR OCCACCOCGA TOPOLIGACIOU GETTAACGOG CARCATOCCC GTGACGTCAT 1000 CTOGGTGAAC TGGCRARCCA AGTCGGGCGG CACGCGTACA GGGAACGTGA CATTGGCCGA 1140 GGGACCCCCG GCCTGATTTC TCOCGGATAC CACCCGCCGG CCGGCCAATT CCATFGGCGC 1200 CAGOCGIVAT IGOOGGSTGA GOOGGGGGST TOOGTOTOCC STREEGGGGGG CATTONNIAA 1260 GCAATGRAUG REGURGAACA CASCUTTUAG CACCUTCOOG TECRAGOGCAG TYACCUTCULA 1336 ASCISCIVITGS TOGASCATOU SGRISCOMAS SACTIONNUM GOSCOSCOSC COTECCODO: 1380 GATCUGACCT GGTTTAAGCA COCCGTUTTC TACGAGGTGC TOOTCCGGGC GTTCTTCGAC 3440 GCCAGCGCGG ACGGTTCOGN CGRTCTGCGT GURCYCATOS ATCGCCTCGA CTACCYGCAG 1500 TOSCITICA TEGACICAT CICTICOCCC COTTECTACS ACTUROCOCT GEOGRAPHIC 1560 GGTTACGACA TTOGOGACTI CTACAAGGTG CTGCCCGAAT TOGGCACCGT CGACGATTIC 1625 STORUCTURE TURACHOUS TOROUGHUSA COTATORICA TEATCACUEA COTOGTÉATE 1880 BATCACACCT COMMUTOSCA COCCESCITT CANGAGICCO GCCGCGACCC AGACGCACCG 3360 TACONTORCY ATTACGISTS GASCARCACC ACCRACGCT ACACCARGE COCKARGES 1800 TECCTEGACA COGRAGACTO GARCTOSTON TYCCATOOTS TOOGCORAGE CTENCYACTG 1880 GCACCGATTC TT 1872

74

(1) BEQUENCE CHARACTERISTICS:

(A) LENGTH: 1482 base pairs

(W) TYPE: nucleic sold (C) STUAMDEDNESS: single

(B) TOPOLOSY: linear

(xi) SEQUENCE DESCRIPTION: SEC TO NO.18:

CTTCGOCGAA ACCTGATGOC GAGGAACAGS STW7TOCCGT GAGGCCGACG GCGTCCGACC 60 CCGCCCTCCT CCCCCAGATC AGGCAGTCGC TTGATGCGAC AAAACGCTTG BCCAGCGTGC 120 ACCTAGOOGT COMAACAACC GGGAANSTUG ACAGCTISCT GGGTATTWCC ACTGCCGATG 180 TOGACOTOCS SGOCAATCCO CIOSCOSCAA AGGCCSTAIG CACCIACAAC GACGAGRAGA 240 GTGTCCCGTT TCGGGTACAA GGCGACAACA TCTCGGTOAA ACTGTTCGAC GACTGGACKIA 300 ATCTCGGCTC GATTTCTGAA CTGTCAACTT CACGCGTGCT CGATCCTGCC GCTCGGGTGA 360 COCROCTOCT STOCKSTSTC ACGAACCTCC AAGCGCAAGG FACCGAASTG ATAGACCGAA 676 TTTOGACCAC CARACTCACC GOGACCATCO COCCGAGOTO POTCAAGATA CTTGATCCTO 480 SCOCCARGAG TECANORIUS CINNOCUTET GENTTECCCA GENEGACITES CACCACCTEG 540 TODGAGOGAG CATOGACCTO GGATOCOGGT CGATTCAGCT CACGCACTOG AGATCHAACG 600 ASCUCSTUAR COTOGRATAC CONGRAGATE COTOGRAGOS TESCECORRA COCKETTORS 660 AACOSTOTCA ACCOCACCCG ANANCTSACC COCTGACGC ATCTGAAAAT TCACCCCCTA GACCAGACGA TYGGYGGYTA TYCTYCGGYG GYYCCGGCYG AMAGGAGGGG OCHGAGAYCG 24.6 COUNCITY OF GOODSTANCE GROCETTES ACCORDANCE CITCAGONES CHOKARCAGA 840 OGGTOGRATOR TOGGOGGCAGO ARCGROSTOG TOGGOGGCCCA ARROTTOGGC COROCGGGGG 200 RAGGOOFFAT TEGACGIGAC GATCARGOIS OCCOGOTORY ACCOGGROUN CACCAGOING 365 AAGAAGAGGT TEGESEGETE GEGETEAAAC GGAATGTANE CCACTTEGTE AAGUACGAGG 1020 AGGMEATAGE SECENAACES SETSACTTES COSTACATES DECEGOOGTS GTGAGGCTES 1080 GOGABOOSTS CTROCCATEC OGCOGOSSTS SOGRACIAGEA COMMATGROS OGCOTURORO 1140 SUSCUTATOS CURSTOCERO ESCARGATOR STETTOCORG ISCURGODOS CUECARRAR 1200 CACGACGITA TOROGGOCOG INATHANATO CAGROTOCCO AGATETICOA TOXTOTOGGO 2260 TITGAGGCCA CGAGCATGCT CRAAGTCGAA CTCTTCCAAC GACTTCCGAA CYCCYCAACTC 1320 SGCCGCGCGG ATGCGGCCCT CRCCRCCATS GCACTCCCCG CCTGACACTT CCCGCTGCAS 1380 SCASSCERCE AGGRAFFETT CONGRETICA GIFFETCGGGG CAGGCGCGAT CHACTAGTCG 1440 SGACACTGAC TCACGCAGGG TGGGGGGCTTT CANTGCTCTT GT 1482

PCT/US97/18293

75

| INFORMATION | | | |
|-------------|--|--|--|
| | | | |

| 123 | SHOU | MAKE 1 | EAB | ACTE | RY | 971 | CS: | |
|-----|------|--------|-----|-------|----|-----|-----|---|
| | | LENGT | | | | | | 2 |
| | (3) | TYPE. | 233 | c.Lo. | C. | aci | d | |

(C) STRANDEDNESC: single (G) TOPOLOGY: lineer

(D) TOPOLOGY: Timeer

(xi) SEQUENCE DESCRIPTION: SEQ IS NO:19:

| GAATTCGGCA | CONSCRIGORS | ATAGCTTCTG | SCOOLGOOD | GACCAGATOG | CICGAGGGTT | 60 |
|------------|-------------|-------------|--------------|-------------|-------------|-----|
| COTOCYCAGG | GCCACCGCCG | GOODCACCAC | CCTRACERRIT | CAGGGCCTGC | AACACGCOGA | 120 |
| CORTCACTOR | TYGCTGCTGG | ACCCUACCAA | cccccccctc | GTTGCTTACG | ACCCGCCTT | 180 |
| OGCCTACGAA | ATCGGCTACA | TCGRQGAAAC | CGGACTSOCO | AGUATOTOCG | GGGAGAACCC | 240 |
| OTAGAAGATO | TECTTCTACA | TOACCSTCTA | CAACGAGCCG | TACGTGCAGG | COCCGGAGCC | 300 |
| GGAGAACTTC | GATOCOGAGG | COUTOCTOO | GGGTATCTAC | CONTATCAGG | CBGCCACCGA | 360 |
| GCAACGCACC | AACAAGGNGC | AGATCOTGGC | CTCCGGGGTA | GCGATGODOG | CGGCGCTOCG | 420 |
| GGCAGCACAG | ATGGT@GCGG | CONSTRUCT | TRICSUCGOO | SACOTOTOGT | COSTGACCAG | 485 |
| TTGGGGCGAG | CTARACCOCK | ACGGGGTGGT | CATCGAGACC | COTTONADA | SECULATOURS | 540 |
| resseesace | OGCGTCCCCT | ACOTGACGAG | AGCCCTTCGAC | AATGCTCGGG | GCCCCGTGAT | 600 |
| coccerence | GACTGGATGC | GCCCCCCACCC | CGAGCAGATC | CGACOSTGGG | TOCOGGGCAC | 660 |
| ATACCTCACG | TTGGGCACCG | ACCOMPTTON | TITTTTCCOM | ACTOSGÓCICA | cosercercs | 720 |
| TYACTTCANC | ACCGACGCCG | AATCCCAGGT | TOSTICOCOCCI | TTTGGGAGKK | GYTGGGGGGG | 780 |
| PUBACEGETE | AATATCGACC | CATTCGGTGC | COSTUSTOSS | CONCCOSCO | NITTACCCGG | 840 |
| STTEGACGAA | GGTGGGGGGGT | TGOSCOCGAN | TAAGYT | | | 876 |
| | | | | | | |

(2) INFORMATION FOR SEQ 10 NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1021 Base pairs
 (A) TYPE: Duclair and
 - (C) STRANDEDNESS: single (S) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: BED 1D NO:20:

AFCCCCCCC CCTCCRAGAR TTCGCCRCGA GAGACAAANT TCCACCCTT AATSCAGGAR 60
CAGATTGATA ACGAATTCAC ACCCCACAA CAATATGTCC CCGTCCCCCT TATTTCGAC 128

| AGCGAAGACC | TOCCOCASTT | GGCGAAGCAT | TTTTACAGCC | ANOCOGECCA | OGAACGAAAC | 188 |
|------------|------------|------------|-------------|--------------|-------------|-------|
| Catguarya | TECTOSTOCA | ACACCTGCTC | GACCGCGACC | TTCKTGTCGA | AATTOCCGGC | 246 |
| GTAGACACGÓ | TOCGRANCIA | GTTCGACAGA | CCCCGCGAGG | CACTOGCGCE | GGCGCTCGAT | 300 |
| CAGGAACGCA | CASTCACCGA | COAGGTCGGT | CEGOTOACAG | CONTROCCCO | CGACGAGGGC | 360 |
| GATTTOCTÉG | GUGAGCAGTY | CATGUAGTGO | TECHTONNAS | AACAGATOSA | AGAGGTGGCC | 4.20 |
| TTGATGGCAA | cccreerece | GGTTGCCGAT | 02000000000 | CCAACCTGTT | CGACCTAGAG | 430 |
| AACTTEGTOG | CACCTGAAGT | GGATGTGGGG | CCOGCCGCAT | caesococcocc | GCACGGSGCC | 546 |
| egeegccecc | TCTAGATOCC | TGGGGGGGAT | CAGCGAGTEG | TOCOGTIFOR: | COGCCCGTCT | 600 |
| TCCAGCCAGG | COTTROTOCS | GCCGGGGTGG | TOAGTACCAR | TCCAGGCCAC | CCCGACCTCC | 660 |
| CGGNAAAAGT | CGAFGTCCTC | STACTCATCS | ACCTYCCAGG | AGTACACCGC | COGGCOUTGA | 726 |
| GCTGCCGAGC | GGTCNACGAG | TECCOGATAT | FOOTTWANG | CAGGCAGTCA | GGGTCCCACG | 780 |
| GCGGTTGGCC | CRACEGGGGT | GGCCGCACTG | CTOOTCAGGT | ATOSGGGGGT | CTTEGGGAGG | 849 |
| AACAACGTOG | GCAGGA69GG | TGGAGCCCGC | ORGATOUGOA | CACCCGGGGG | GCGAAAACGA | (131) |
| CATCAACACC | GCACGGGATC | GATCTGCGGA | GGGGGGTGCG | CKRATACCGA | ACCOCTOTAG | 960 |
| GAGCGCCAGC | AGTTGTTTT | CCACCAGCGA | AGCOTTTTCO | GOTCATOGOS | GOCKNITTAAG | 1020 |
| r | | | | | | 1021 |
| | | | | | | |

(2) INFORMATION FOR SEC TO NO. 21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 321 bass pairs
- (B) TYPE: nucleic acid (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(RI) REDURNCE DESCRIPTION: SEG ID NO.21:

OGTGOOGADG AACGGAAGAA CACAACCATG AAGATGGTGA AATGGATCGC OSCAGGTCTG 60 ACCECCUCOG CTOCAATOUG CERCECTEGG GOOGGTGTWA CTTCGATCAT GECTGGOGGR COGGTOSTAT ACCAGATGCA GEOGGTOSTO TECRGUEGO CACTROCETT DESCUEDANA 180 TERGESTERTS AMPRICEDAD COCCOCCOMO TEGESCACAD SOTICACAD SOTICACAD 240 COCAACGIGI CETTIFONGAA CAAGGINAGT CITOTICGAGG GNGGMATCEG NGCHAMCGAG 330 CONGRGNATE CHEGANCACA A 322

(2) INFURMATION FOR SEC ID NO.32:

(i) SEQUENCE CHARACTERISTICS:

27

| (8) | LENGTH: | 373 | basse | 033 61 |
|-------|--------------|---------|-------------|--------|
| (8) | TYPE: n: | oclea | c aci | d |
| 603 | STRANDE | MESS | : 830 | gle |
| 0.000 | 28-70-YE 000 | 2 . 3 3 | ASSESSED BY | |

(RI) SEQUENCE DESCRIPTION: SEG 15 NO:22:

| TCFFATCGGT | TOOGGTTGGC | GACGGGTTTT | GEOMETRISCI | GGTTARCOON | CPCGGCCAGC | 8 |
|-------------|-------------|------------|-------------|-------------|------------|-----|
| OGATOGACGG | GCGCGCAGAC | GTOGACTOCS | ATÁCTOGGCG | CGCGCYGGAG | CTOCAGGGGG | 12 |
| COTOGGTGGT | GNACCGGCAA | GGCCTCAAGG | AGCCCTTONA | GACCGGGGATC | AREGUCATTO | 2.8 |
| ACGCGATGAC | COCSATOGGC | OGCOGGCAGC | GCCAGCTGAT | CATOGGGGAC | CGCAACACYC | 24 |
| SCAAAAACCG | contendiate | COCACACCAT | COTCAAACGA | GCGCGAAGAA | CTGGGAGTCC | 36 |
| GGTCKGATCCC | AAGAAGCAGS | TOCCCTTCTC | TATACGTTGG | CCATCGGGCA | AGAAGGGGAA | 36 |
| CTTACCATOS | cos | | | | | 3.7 |
| | | | | | | |

(2) INFORMATION FOR SEC 10 W0:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTS: 352 base pairs

thi TYPE: medain soid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(RI) SEQUENCE DESCRIPTION: SEG ID #0:23:

| GTGACGCCGT | GATGGGGATTC | CTRESCORDE | COGGTCCSCT | SOCCETTORIC | GATCAGCAAC | 61 |
|------------|-------------|-------------|-------------|-------------|------------|-----|
| TOGTTACOCG | GGTOCCGCAA | SECTION | TTGCTCAGGC | #GCOGCTGTG | COGTGGTGT | 120 |
| TOTIGACGGC | CTGGTACOGG | TTGGCCGATT | TACCCGAGAT | CAACCUGGEC | GAATCGGTQC | 180 |
| TGATCCATGC | COGTACCOCC | GGTGTGGGCA | TSSUGGETTER | GCAGCTGGCT | CCCCATTOGG | 280 |
| OCCTOGAGCT | TTTCGFCACC | SCCAGCCGTG | GRAAGTGGGA | CACCOTGCTC | GCCAINGNGT | 300 |
| TIGACGACGA | DCCATATOGO | NGAST CCCSC | ACATNOGAAG | TTCCGANGGA | GA | 362 |
| | | | | | | |

(2) INFORMATION FOR SEC ID NG:24:

- (1) SEQUENCE CHARACTERESTICS: (A) LENOTE: 726 base pairs (9) TYPE: nucleic soid (C) STRANDEDNEED: sangle
 - (S) TOPOLOGY: linear

PCT/US97/18293

| | BESTRIPTION | | |
|--|-------------|--|--|
| | | | |

| GAAATCCGCG | TTEATFOCGT | TOGACCAGCG | QCTQCC5ATA | ATOGAGGAAG | TGATCABCCC | 59 |
|-------------|-------------|------------|-------------|------------|------------|-----|
| GOGGTTCUCS | GCGCTCATGG | GTCACAGOGA | GTAATCAGCA | ASTICTCIOS | TATATCGCAC | 120 |
| CTAGCGTCCA | GITGCTTGGC | AGRTOGOTTT | COPACOGTCA | TOGGATGTAG | COUNTGECOT | 190 |
| 000000XCGUT | CATACTERCO | GOSTGCATCC | 10000040000 | TOTGGCGGGT | CTCGGGGTCG | 240 |
| OCOGGGAGYC | CGCAGCCCAS | ACCGCGCCGG | TECCCONCTA | CTACTOGYG: | CCCCCCCAGC | 300 |
| CTTTCGACCC | CGCATGGGGG | CCCAACTGGG | ATCCCTACAC | CFGCCATGAC | GACTTCCACC | 369 |
| OCGACAGOGA | OSCCOCORAC | CACAGCCGCC | ACTACICCOGO | ACCCATOUTO | GAMGGTOCCS | 420 |
| TGCTTGACGA | TOCCOGGIGGT | GCGCCGCCGC | cccesserse | CGGPGGCGGC | GCATAGCGCT | 480 |
| CETTGAOCOG | OCCOCATCAG | CGAATACGCG | TATABACCICG | 990STGCCCC | COGCAAGCTA | 540 |
| CEACCCCCCG | CGGGGGAGAT | TTACGCTOOC | GTGCCGATGG | ATCGCGCCGT | COSATGACAG | 600 |
| AAAATAGGCG | ACCOTTTTGG | CAACCOCTTG | GAGGACGCTT | GRAGGGARCO | TGTCATGAAC | 660 |
| GGCGACASCG | CCYCCACCAT | CGACATCGAC | AAGGTTGTTA | CCCGCACACC | CETTOGCOGG | 720 |
| ATCGTG | | | | | | 726 |
| | | | | | | |

(2) INFORMATION FOR SEQ TO NO.25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 580 base pairs

(B) TYPE: mucleic soid (C) STRANGEDMESS: single

(D) TOFOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

OFCINCUACU ACGRACETOS GEOCEACORO COCCERTOOS TEGATOCARS CONCENCIONAS 66 SCTOSCOGAC CATATORNAS CATSCIGGET SCCONCIGAS CGACCITITS ACCASCIGAS 120 CTGCCCGATG OXGGCCCGGT GAAGTCATTG COCCGGGCCT TGTGCACCTG ATGAACCCCA 1.85 ATMOGGAACA RYAGGGGGGT GAFTTGGCAG TECRATOFCG GGTATGGCTG GAARTCGART 240 GECEGGECAT GCTCEXXXIC CACCAGNOTIC GCGCAGGCGG GCCAGGCGGA ATCTGGAGGG 300 AGCACTCANT GEODECKNIE AAGCCCCCGA COSSCGACGG TCCTTTGGAA GCAACTAAGG 3.60 ASSESSED CATTOTORIC CORSTACCAS TECROSOTOS DESTCUCCOS SICETOGRAS. 120 TGACACCOGA CGAAGCCGCC GCACTGGGTG ACGAACTCAA AGGCGTTACT AGCTRAGACC 180 AGCCCANCOS CHARTGETOS SOSTFACGOS SACACCTTOS GETAGATESTO CACESTATAS 540 TOGGCGATGT ATGCCCAGGA GAACTCTTGG ATACASCECT 590

| GETACOSCO SETTOTOS TETCOSCOS GECONTOSCO COGGAGACAA COCCATOGOC GETOTICACGO STACETOSCO CADCACACCO GETOCATOCO [3] INFORMATION FOR SEQ ID NO:27: (A) SENDENCE CHARACTERISTICS: (A) LEMOTH: 272 base pairs (A) LEMOTH: 272 base pairs (A) LEMOTH: 272 base pairs (A) CENTRADELNESS: single (D) TOPOLUSY: Linear [X1] SEQUENCE DESCRIPTION: SEQ ID NO:27: GACACCHATA CHATACTICAT UTACGCCDAC STTOTOCCA GECOGGTECT AAGGCGATG GAATGACAA GECTACTOCC GAATGACCC CHICGTCGCT GAACGCGAAC ARTOGGGAA COGCAACAA SECTIOSCOT TOCCCAAGACC CHICGTCGCT GAACGCGAAC ARTOGGGAAC ACCCCAACAA SECTIOSCOT TOCCCAAGACC CHICGTCGCT (A) LEMOTH: 317 base pairs (B) TYPE: DUDICAC SINGle (C) STAMBELENESS: Single (D) TOPOLICY: Libear [X1] BEQUENCE CHARACTERICS: (A) LEMOTH: 317 base pairs (B) TYPE: DUDICAC SINGle (C) STAMBELENESS: Single (D) TOPOLICY: Libear | | |
|--|--|-----|
| (A) LENGTH: 106 base pairs (B) TYPE: Nucleis acid (C) STRAMBERNASS: Mingle (D) TOFOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26: AACEGAGGCG CCGGGGGTTT TGGCGGGGCC GGGGGGTCG GGGGGGGCACGG GGGGGGCCGC GGTACCGCC GGAACGGCC GGAACGGCC GGAACGGCC GGAACGGCC GGAACGGCC GGAACGGCGACCC GGAACGGCGACCC GGAACGGCGACCC GGAACGGCGACCC GGAACGGCGACCC GGAACGGCGACCC GGAACGGCACCC GGAACGGCGACCC GGAACGGCACCC GGAACGGCACCC GGAACGGCACCC GGAACGACCC GGAACGACCC GGAACGACCC GGAACGACCC GGAACGACCC GGAACGACCCC GGAACGACCC GGAACGACCC GGAACGACCC GGAACGACCC GGAACGACCC GGAACGACCCC GGAACGACCC GGAACGACCCC GGAACGACCC GGAACGACCCC GGAACGACCCC GGAACGACCC GGAACGACCC GGAACGACCCC GGAACGACCCC GGAACGACCCC GGAACGACCCC GGAACGACCCC GGAACGACCCC GGAACGACCCC GGAACGACCCC GGAACGACCCC GGAACGACCCCC GGAACGACCCC GGAACGACCCCCCCC | (2) INFORMATION FOR SEQ IO NO.26: | |
| AACGAGGC CGGGGGTTT TGGCGGGGC GGGGGGTCG GGGGGAACGG GGGGGCGC GGGAGGCC GGGAGGCC GGGAGCCAACGGC GGTACACGCC GGTACACGCC GGTACACGCC GGTACACGCC GGTACACGCC GGTACACCCG GGTACACCCGC GGTACACCCGC GGTACACCCGC GGTACACCCGC GGTACACCCGC GGTACACCCGC GGTACACCCGC GGTACACCCGC GGTACACCCGCACCCC GGTACACCCCCC GGTACACCCCCC GGTACACCCCCC GGTACACCCCCC GGTACACCCCCC GGTACACCCCCC GGTACACCCCCC GGTACACCCCC GGCGGCTGCC GGTACACCCCC GGTACACCCCC GGTACACCCCC GGTACACCCCC GGTACACCCC GGTACACCCCC GGTACACCCCCCCCCC | (A) LEMSTH: 100 base pairs (B) TYPE: Nucleic sold (C) STRAMMERMENT: single | |
| GUTACOSCO SETTOTOS TOTOGOGOS GUCOTOGOS CUGGAGACAA COCCATOGOC GUTOTCACOG STACOTOGOS CADCACACCO GUTOCATOCO [3] INFORMATION FOR SEQ ID NO:27: (A) SECUENCE CHARACTERISTICS: (A) LENCTH: 212 base pairs (A) LENCTH: 212 base pairs (B) TYPE: INDIBIS SI INJE (D) TOPOLOGY: LINEAR [X4] SECUENCE DESCRIPTION: SEC ID NO:27: GACACCHATA CHATGHOLOGO GACCATOGO GUTCACGATO 6 CASUBCACAC DUBACOGUET GACCATOGO GATOCACGA GOCGGOTOGO AAGGCGATG GAATGACAA GUTGCGGGTA ATTCATACOG GAATGACCC CHICGICGCT AAGGCGAAC ANTOGOGACA COSCAACAAC ACCTTOGOGO TOTOCTOGOT GAACGGGAC GUAACGTACA GACCATOGO GAATGACCC CHICGICGCT (A) LONGTH: 317 base pairs (A) LONGTH: 317 base pairs (B) TYPE: Bublic scid (C) STAMBERNESS: Single (D) TOPOLOGY: LINEAR [X1] BEQUENCE CESCRIPTION: SEG ID NO:28: (X1) BEQUENCE CESCRIPTION: SEG ID NO:28. | (xi) SECHENCE DESCRIPTION: SEQ IS NO:26; | |
| GETOTCACGG STACGTOGGC CADCACCCG GETOCATCCS [3] INFORMATION FOR SEQ ID NO:27: (A) SECUENCE CHARACTERISTICS: (A) LENGTH: 212 base pairs (A) LENGTH: 212 base pairs (A) LENGTH: 2021 (C) STAMUBLINESS: simple (D) TOPOLUSY: Linear [X4] SECUENCE DESCRIPTION: SEC ID NO:27: GACACCHATA CHATGKTGAT UTACGCCDAC GTTGACGATC (A) CASUBCACAC CHARACTERISTICS GACACCHATA CHATGKTGAC GACACCHATA CACACCHATA CHATGKTGAC GACACCHATA CHATGKTGAC GACACCHATA CACACCHATA CHATGKTGAC GACACCHATA CACACCHATA CHATGKTGAC GACACCHATA CHATGKTGAC GACACCHATA CACACCHATA CHATGACCHATA CACACCHATA CHATGACCHATA CACACCHATA CAC | ANCEGAGETE COGGOGTTTT TEGEGGGGCC GEGGCGGTTG GGGGCAACGG GGGGCCCGCC | 6 |
| (A) SECURNCE CHARACTERISTICS: (A) LENGTH: 272 base pairs (A) LENGTH: 272 base pairs (A) LENGTH: 272 base pairs (B) TYPE: nucleic acid (C) STRANDELNESS: simple (D) TOPOLOGY: Linear (X1: SEQUENCE DESCRIPTION: SEC ID NO:27: GACACCGRATA CHARGNETAT CTACCCCAAC CTTOTAGGC CTTCACGATC CASHDCRICE DOGACGUGT GACCATCGC GATEGRACC COTTCTAGGC CTTCACGATC CASHDCRICE DOGACGUGT GACCATCGC GATEGRACC COTTCTCCCT 18 GACCGGAAC ANTOHACAG COCCAACAAC ACCTTCGCGT TGACCCCORS TGTCATTGTC 24 GACCGGAAC ANTOHACAG GACCATCGC CQ T2) INFORMATION FOR SEG ID NO:28: (A) LENGTH: 171 base pairs (A) LENGTH: 171 base pairs (C) STRANDELNESS: Simple (D) TOPOLOGY: LINEAR (X1: BEQUENCE CESCRIPTION: SEG ID NO:28: (X1: BEQUENCE CESCRIPTION: SEG ID NO:28: | SCHACCECCO SETTETTOS TETCOSCOCO COCCAGECA CECANOCOC | 120 |
| (A) SMODENCE CHARACTERISTICS: (A) LENGTH: 212 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: simple (D) TOPOLOGY: Linear (X1: SEQUENCE DESCRIPTION: SEO ID NO:27: GACACCGRATA CHARGNEST STREEDCARC STTURESGA GCOGGCTECC CASCCGRATA CHARGNEST STREEDCARC STTURESGA GCOGGCTECC 12 RAGGCGATG GAATGRACA GCTGCGGGTA ATTCATACCG GAATGRACCC CUTCHTCGCT 18 GACGCGRAC MATCHGACA GCTGCGGGTA ATTCATACCG GAATGRACCC CUTCHTCGCT 18 GACGCGRAC MATCHGACA GCTGCGGGTA ATTCATACCG GAATGRACCC CUTCHTCGCT 18 GACGCGRAC MATCHGACA GCTGCGGGTA ATTCATACCG GAATGRACCC CUTCHTCGCT 18 GAACGCGRAC GAATGRACA GCTGCGGGTA ATTCATACCG GAATGRACCC CUTCHTCGCT 18 GAACGCGRAC MATCHGACA GCTGCGGGT TGGCGGCCGGG TGTCTTTGTC (A) LENGTH: 317 base pairs (A) LENGTH: 131664 (C) STRANDEZNESG; SINGAR (D) TOPOLOGY: 131664 (E) TITC. (X1) BEQUENCE CESCRIPTION: SEG ID NO:28: GCGGCGGGTG STTCTCGGAC TATCTGCGCA CGGTGGAGGTG CGCGGGCTGA | SETSTCACGE STACSTCGGE CASCACACCG GETEGATCCG | 168 |
| (A) LENGTH: 212 base pairs (B) TYPE: Number of odd (C) STRENDENESS: single (D) TOPOLOGY: Linear (X1: SEQUENCE DESCRIPTION: SEO ID NO:27: GACACCGRATA CHARGEGAT STRENGECT GATTERDRACA CONTOGRAGE STRENGECT GACACCGRATA CHARGEGATA GACACCGRATA GACACCGRATA GACACCGRATA GACACCGRATA GACACCGRATA GACACCGRATA GACACCGRATA GACACCGRATA GACACCGRATA GACACCGCACCGCT 18 GARGEGARG GARTCGRATA GACCACCGC GATTCRACCC CHICGICGCT 18 GARCEGARG GARTCGRATA GACCACCACC CG GARCEGARG COGCARCARA GACTTEROGT TGCCCCCCCC CHICGICGCT 18 GARCEGARG COGCARCARA GACCACCACCC CG (1) SEQUENCE CRARROTERISTICS: (A) LENGTH: 317 base pairs (B) TYPE BEDDIACA COCCACCCC (C) STRANGERAGES SINGLE (C) STRANGERAGES SINGLE (X1: SEQUENCE CESCRIPTION: SEQ ID NO:28: GARCEGARG TTTTTCGRAC TRICTICRICA COCCARCAGE COCCARCAGE GACACCCGGT TTTTTCGRAC TRICTICRICA COCCARCAGE COCCARCAGE GACACCCGGT TTTTTCGRAC TRICTICRICA COCCARCAGE COCCARCAGE GACACCCGGT TTTTTCGRAC TRICTICRICA COCCARCAGE COCCARCAGE GACACCCGT TTTTTCGRAC TRICTICRICA COCCARCAGE COCCARCAGE GCACCCGGT TTTTTCGRAC TRICTICRICA COCCARCAGE GCCGGCGGCG COCCARCAGE GCACCCGGT TTTTTCGRAC TRICTICRICA COCCARCAGE GCCGGCGGCGT COCCARCAGE GCACCCGGT TTTTTCGRAC TRICTICRICA COCCARCAGE GCCGGCGGCGC COCCARCAGE GCACCCGGT TTTTTCGRAC TRICTICRICA COCCARCAGE GCCGGCGGCGCG CCCCARCAGE GCACCCGGT TTTTTCGRAC TRICTICRICA COCCARCAGE GCCGGCGGCGC CCCCARCAGE GCCGCCGCT TTTTTCGRAC TRICTICRICA COCCARCAGE GCCGGCGGCGCG CCCCACCAGCT CCCCACCACCT CCCCACCACCT CCCCACCT C | [7] INFORMATION FOR SEQ IS NO:27: | |
| GAGACGRATA DIATRATGAT STACRODAR STTUTCHACA CECTURAGE STTUACRATC CASEBORATE DEBACROUST GAUCATORE GATSCHROOD STTUKEGA GEOGRETECE RAGGOGRATE GAATCACACA SCTEGOSGSTA ATTCATACOG GAATGACCC CUICGTCGCT 18 GARGGOGRAD ANTENGACACA COSCANCARA ACOTTEGOST TECHNOCOGS TETCETTOTO 20 BOUTAGGAGU GUAACCTACA GACDARCOCC CG 27 (1) INFORMATION FOR SEG ID BO:28: (1) SEQUENCE CHARACTERISTICS: (A) LEMKTR: 317 base pairs (B) TYPE HODIALS OLD (C) STRANBERNESS: SINGLA (D) TOPOLOGY: LISBORY (X1) BEQUENCE CESCRIPTION: SEG ID BO:28: GCAGOCGGTG STTUTCGRAC TATOTGCRA COGTENTSCA GCGCGAGGTG COCCAGCTER 6: | (A) LEWIH: 272 base pairs (B) TYPE: nucleic soid (C) STREMDENSES: single | |
| CASEDCARAC ECGACGUST GACCATCREC GATSCRECC CSTTCKEGA GOGGETECC 12 RAGGCGATGG GAATCGACAA GCTGCGGGTA ATTCATACCG GAATGGACCC CSTCGCTGCT 18 GAACGGAAC ASTCBGACGA COGCAACAAC ACGTTGGGT TGCCGCCGG TGTCGTTGTC 2c BECTAGGACC GGAACGTACA GACCAACCCC CG [2] INFORMATION FOR SEG ID BO:20: [1] SECUENCE CHARACTERISTICS: [A] LERKTH: 317 base pairs [A] LERKTH: 317 base pairs [C] STRANBECKRESS: Single [D] TOPOLOGY: Lisbeat [X1] BEQUENCE CESCRIPTION: SEG ID BO:20: GCAGCCGGTG TTCTCGGAC TATCTGCCCA CGGTGACGTG CGCGAGCTGA 6- | | 47 |
| AAGGCGATGG GAAYCGACAA GCTGGGGGTA ATTCATACOG GAATGGACCC CUTCGTCGCT GARGGGGAC AUTOBGACGA COGCAACAAC ACGTTGGGT TGCCGCCGGG TGTCGTTGTC SCCTAGGAGC GUAACGTACA GACCAACCCC CG [2] INFORMATION FOR SEG ID BO:20: (i) SECUENCE CHARACTERISTICS: (A) LERKTH: 317 base pairs (A) LERKTH: 317 base pairs (B) TYPE unpoles ocid (C) STRANGECORGS: single (D) TOPOLOGY: Linear [Xi) SEQUENCE CESCRIPTION: SEG ID NO:20: GCAGCOGGTG STTCTCGGAC TATCTGCDCA CGGTGATGGG CGCGAGCTGA GCGCGAGCTGA | | |
| GRACOGGARC AUTORGACCA COGGARCARA REGITORGET IGRIGORGE INFORTUTO 26 BUCTAGGAGG GRACOTROA GREDARCUCC CG [2] INFORMATION FOR SEG ID NO:28. [1] SEQUENCE CHARACTERISTICS: [A] LEMCTH: 317 base pairs [A] LEMCTH: 100 beach old (C) STRANBERNESS: Simple [D] TOPOLOGY: Linear [X1] BEQUENCE CESCRIPTION: SEG ID NO:28: GRACOGGAGG STRUCKSAC TATOTOCHA CONTENTSCA GUGGGAGGIS COCCAGOTRA | | |
| SUCTAGENCE GEARCHTECH EMERANCICE OF 12) INFORMATION FOR SEG ID NO.28: (i) SEQUENCE CHARACTERISTICS: (A) LEMKTH: 317 heas pairs (E) TYPE: INDIAN OLD CHE (C) STRANSECONESS: Simple (D) TOPOLOGY: Linear (X1) DEQUENCE CESCRIPTION: SEG ID NO.28: GEAGLOGGET STETECHAR CHARGEOGRA GEGEGRAGIE COCCARCTER 6: | | |
| (2) INFORMATION FOR SEG ID NO:22: (3) SEQUENCE CHARACTERISTICS; (4) LEMTH: 317 hear pairs (5) TYPE models acid (6) STRANGERMENS: single (D) TOPOLOGY: Linear (xi) BEQUENCE DESCRIPTION: SEG ID NO:28; GORGOOGGETS TITCTORGE TATOTOCOME COGRACTER 6: | | 272 |
| (i) SECUENCE CHARACTERISTICS: (A) LOWETH: 317 base pairs (B) TYPE: uncloid solid (C) STRANGEOWERS: slugle (D) TOPOLOGY: linear [xi] REQUENCE DESCRIPTION: DEC ID NO:28: GRECORGET STRUCTURES CONTENDED GRECORGETS COCKAGCTBA 6: | | |
| GC860CGGTS STTCTCGGAC TATCTGCGCA CGGTGACGTCA GCGCGACGTG CGCCGAGCTGA | (A) LEMCTH: 317 here pairs (B) TYPH: muchaic scied (C) STRAMMENNES: simple | |
| | (xi) BEQUENCE DESCRIPTION: SEQ 10 NO:28; | |
| ASCOGNICOS GUAGACORAT COULTGUES GATTUATECE CYACUTERCO SQUATEACCO 12 | GCRGCCGGTG STTCTCGGAC TATCTSCRCA CGGTGACGCA GCGCGACGTT CGCGAGCTGA | 60 |
| | ASCSCATOGA SCAGACGRAT CSUCTSCOSC SCTTCATGCC CTACCTSNCC SCTATCACCG | 120 |

| COCADBACCT GRACGIFGCC GRAGGEGGCG GCGCATCGG GCT/CACCGC GGCACGATCC | 100 |
|---|-----|
| STICEGRATOR SECURETIES SACACCORDE ARCHIGORACA TOSCOTOCOC GCONSCIDOS | 240 |
| GGARTCHGAC CGCGAAGATC RAGAAGCOOF CAAAGATCCA CSTCGTCGAC ACTGGUTTUG | 300 |
| COCCAROLLA OCCUROR | 317 |
| (2) INFORMATION FOR SEC ID NO.29: | |
| (1) BEQUENCE CHARACTERISTICS: (A) LINCTH: 182 base pairs (B) TYED: naclaic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear | |
| (*1) SEQUENCE DESCRIPTION: SEQ ID NG. 29: | |
| SATEGYÉGAG CTGTEGATGA ACAGCGTTGE COGACGCGCG GCGGCCAGCA CGTCGCTGTA | 68 |
| GCAMOSCOGE ACCROCTOGO CEGTEGECAS CATGETEATS ACCROCTOGG COTOGGCCAC | 120 |
| CECTTCEGGC GCGCTACGAA ACAUCGCGAC ACCGTECGCG GCGGCGECCG ACCCCGCCGT | 185 |
| 66 | 195 |
| (2) INFORMATION FOR USO TO NO:30: | |
| (1) SEQUENCE CRAFACTERISTICS: (A) LENGTH: DOE Hass pairs (B) TPED: Invalid's would (G) STRONDEDNESS: single (G) TOPOLOGY: linear | |
| (xi) SEQUENCE DESCRIPTION: SEQ 10 NO:30: | |
| GATOROGRAG TITGOTORGO AGGTGGTOGA OGRGRAGOTO TGRUNDEUTG OGRAGORGOT | 60 |
| GGGCGTTCAC GAGGCGAAGA CACGCCTGTC CGAGCTGCTG CGGCTCGTCT ACGGCGGGCA | 120 |
| GAUGTTHAGA TTGCCCGCCG CDXCGAGCCG GTAGCAAAGC TTGTGCCGCT GCATCCTCAT | 180 |
| SMEACHCEGG GETTRESCAT TOACCATEGG STETACCEGG TECCCEGADER TITGEACCET | 246 |
| COGFIGICAG ACQUECTUCT CGARCUCTT CALCOUTGAN GCGCTROCTC ATCUMENCOC | 300 |
| ACCTTING | 308 |
| (2) INFORMATION FOR SEQ ID NO.31: (1) SEQUENCE CHAPACTERISTICS: (3) I LENGTH: Z67 Dase pairs (5) ITEE: nucleic acid (1) STRANUMERS: airpais | |

81

(D) TOPOLOGY: Jinear

| (*1: | GEOGRESON | DESCRIPTION: | 2000 C | 7.35 | 100 - B 2 - |
|------|-----------|--------------|--------|------|-------------|

| CCGACGACGA | GCAACTCACG | TOGATGATGG | TOGGCAGOSS | CATTGAGGAC | GGAGAGASTC | 64 |
|------------|------------|------------|------------|------------|------------|------|
| CBOCCGAAGC | TGCGGCGGGG | CAAGTOCTCA | TAGTGACCGG | COSTAGREGG | CTCCCCCGAT | 120 |
| GECACCGGAC | TATTCTGGTG | POCCECTORC | COSTANGASC | GGGTAAAAAA | ATGTGAGGGG | 1.00 |
| ACACGATGAG | CAATCACACC | TACOSAGTGA | TOGAGATOGT | COSGACCTCG | CCCGACGGCG | 246 |
| TOGACGCGGC | AATCCAGGC | GGTCTGG | | | | 267 |
| | | | | | | |

(2) INFORMATION FOR SEQ ID MG:32:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1539 base pairs (B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(P) TOPOLOGY: linear

(ii) and manner warmen

(x1) SEQUENCE DESCRIPTION: SEQ ID NO/32:

CTESTSCOSA AAGARTSTER CROGACAGRA TGAGCASTCA CACCTROCOR STORTEGAGE 86 TOUTOGGGAD CTOGGCCCGAC GGCGTCGACG CGGCATCCA GGGGGTCTG GCCCGACCTA 128 COCAGACCAT SUSCECTE GACTESTICS ARTROPORTY ARTROPASSO CACCISSING 190 ANSGAGUGGT UGUSUACTIC GAGGIGACTA IGAAAGIGGO CIICOGCIGG AGGAITCCIG 240 AACCTTCAAG CGCGGULGAT AACTGAGGTG CATCATTAAG CGACTTTTCC AGAACATCCT 300 GAOGODOTOS ARACGOGOTI CAGOGGACGG IGGOTOCGCO GAGGGCCOTGC CYCCARRATE 365 COTOCCACAA T70010900G GUGCCTACAA GGAAGTUGET GCTGAATTOG T0500TATCT 420 GGTCGACCTC TGTGGCCYCC AGCCGGACGA AGCCGGTGCTC CACGTCAGCT CCACCYCGGA 480 SCSGATGSCS TTGCCGCTCA CCSGCTATCF GRACAGCGAG GGACGCTAGG COGGCTTCGA 540 TATOTOGCAG ARAGOCATOS OSTIGITGOCA EGAGCACATO ACOTOGOGO ACOTORACET 200 OCACTICUAG GICIOGACA IUTACAACIC GOTUFACAAC COGAAAGUUA AATACCADIC 660 ACTRIACTIT OSCITICAT ATOUGGAIGO GIOGITEGAT GIOGITETTO TRACCIONAL CTICACCORC ATSTITCOSC CSGACGISGA SCROTATOTS GACGAGATOT COCCEGISCT 780 GAAGCOCGGC GRACGATGCC TOTGCACGTA CYTCTTGCTC AATGACGAGT CATTAGCCCA 840 CATOUCUGAA GUAARBARTO OGUALDAUTT CYAGUATERO EGACUEGUTT ATCOGRUAT 300

82

| 0 | CACAAGAAG | CORCCOSAAS | MAGCAATOGG | CTTREEGGAS | RECTTURECA | GERTSYCTA | 990 |
|----|-----------|-------------|-------------|------------|-------------|------------|--------|
| T | GGCRAGTTC | GGCCFCGCCG | TOCACGAACC | ATTGCACTAC | GGCTCATGGA | STOSCOGGGA | 1026 |
| A | CCACGCCTA | AGCTTCCAGG | ACATOGROAT | COTCACCAAA | ACCCCCGACCT | ACCTCOCCAT | 1098 |
| C | COGGAACCA | TESCGACACC | STOODSCOUR | GEOCCOCTGC | CGRCAGGETS | ATTAGGOGGG | 1146 |
| 0 | AGATTAGOT | cecescoses | COOGGCTCCA | AGTACGGCOC | CCCXXXXTGGC | GTCACCGGCT | 1500 |
| G | STAACCACQ | CTTGGGGGGG | T0000000000 | CUTGRINGAT | CAGGTOGTAG | ATGCCGRCAA | 1260 |
| À | scerecers | ATCGGTCATC | ACCAACGG16 | ACAGCAGCOG | OTTOTGCACC | ASCGCGAACU | 1,3,20 |
| O | CACCOCCC | CICCOGGICI | GTOCAGCOGA | TEGAGEORGE | CAAGCCCNCA | TGACCAAACC | 1380 |
| (3 | OSSCATCAC | GETCOCCATAC | SOCATACCGT | GATAGOCAAG | ATGRAAATTT | AMENGEACCA | 1400 |
| N. | PAGATTYUG | ATCCGGCAGA | ACTTGCCGTC | SCITGOGGGT | CARGCCCGTG | ACCAGCTOCC | 1500 |
| G | XACAAGAA | COSTATOCOS | TUGATOTOGO | CTOSTGCCG | | | 1533 |
| | | | | | | | |

(2) INFORMATION FOR SEQ TO NO.33:

(i) SEQUENCE CHARACTERISTICS: (A) DENGTH: GSI base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single (B) TOPOLOGY: linear

(MA) SEQUENCE DESCRIPTION: SEQ IS NO. 33:

CEGURGGETG GUSTGENINA SUNTUACOUC GUSTURGOUG GAGUTGACCE CUSTCUCAGUT 80 COGGOTTECT COCCCCCCC ROSACACGGC CTATGGGCTG ROSGTGCCCC CGCCCGGTGAT 1.20 OBCCGAGASC COTOCTUANC TGATGSTFCT GATAGOGNOS AACCTCTTGG GOCAALACRC 189 CXXGGCGATC SCUSTCARCS AGGCCGARTA CGGCGGAGATS TGGGCCCARG ACGCCRACCCC 240 GATGITIGGS TACGOCGCGG CGACGGCCAC GGCGACGING ACGITICCIGC CHITICITAGINA 300 SOCCOCCEAS ATGACCASON CONCTESONY ONTOSASCAS SOCCOCCOCCA TOGACGASON 360 CTCCCACACA GCCGCCGA ACCACTTGAT GAACAATGFC CCCCAGGCGC TGAAACAGTT 426 SOCOCAGEOE ACCCAGGOCA CCACGCCTTC TICCAAGCTG GOTGGCCTGI CCAACACGGIT 980 CTCSCCGCAT CONTOGCCGA TCAGCAACAT CUFCTCCATG CCCAACAACT ACATGTCGAT 244 CACCAACTON OCTOTOTOGA TOACCAACAC CTTGACCTCG ATGTTGAAGO SCITTGCTCC 500 GOOGGOGGOO GOOCAGGOON THINAACTOR GOOGCAAAAU GUGOTOOGGE COATGAGCTC 680 OCTGGGCASC TOGCTGGGTT CTTCGGGTCT GGGCGCTGGG GTGGCCGCCA ACTTGGGTCG 720 COCOGCOTOG STACOGTATS STONOCOGGA TOSCHGARAA TATGCAMAST CTOGTCGOCG 780

| WO 98/16646 | PCT/US97/18293 |
|-------------|----------------|
| | |

| GAAUGGTGGT CCGCCCTAAG GTTTACCCCC GTTTTCTGGA TGCGGTGAAC TTCGTCAAGG | 840 |
|---|-----|
| CAARCAGTTA C | 851 |
| (2) INPORMATION FOR SEQ ID NO:34: | |
| (i) SEQUENCE CHARACTERISTICS: (A) LEMBTH 25% base pairs (B) FYPE: modele acid (C) STRANGEMESS: single (D) TOPOLOGY: linear | |
| (%%) SEQUENCE RESCRIPTION; SEG ED NO(34: | |
| GATGGRICGG GCGGRARITT GGRCCRGATT CGCCTCCGGC GATAACCCAR TCRATCGAAC | 60 |
| CTAGATTTAT TOUGTCOASE GECCUSAGTA ATBOUTUSCA SGAGAGGGAAC CTRACTOUTS | 120 |
| CEGECACITE TOSTAGETEC TEGATACINE GEARGECOTE GACAFTITES ACCGACACEC | 180 |
| CONTREARAC STICGAGGC CASTICUAGUI TGIGAGCGAG GCCACGCAGI CGCAGGCTGC | 240 |
| SCTTOUTCAA GATC | 254 |
| (7) INFORMATION FOR SUQ ID NO.35: | |
| (i) SECTIONIE CHARACTERISTICS: (ii) LEMONTH 1227 bare pairs (iii) TYPE: nucleic acid (c) STRANDOMESS: single (D) TOPOLOGY: linear | |
| (xi) SEQUENCE DESCRIPTION: SEQ 10 NO:35: | |
| GRICOTGAGO GAAGOGGOG COGCCAAGGO GAAGTOGGTU TIGBACCAGG AGGGACGGGA | 60 |
| CORTOTOGCO CTCCGGATCO COGTTCAGCO GCGGGGGTOC OCTGGATTGC SCIAIAACCI | 120 |
| FITTOFFCGAC GACCGGACGC ISGAIGGIGA CCAAACCGCG GAGITCGGTG CTGTCAGGTT | 190 |
| GATORIGGAC CEGATGROOD COCCETRICT GGRAGDCUCG TCGATCGATT TOTTOGACAC | 245 |
| TATEGREARG CARGETTCAC CARCERCART COCCARCECON CUGGETCCTG DECETCCGGG | 300 |
| CATFOSTICA ACTORITARA COCTAGIACO ACCCCOCGO GOGGARGACO TACGAGCAGA | 360 |
| CCANGACCES ACCOCCETEG ANAGCANCE GNECONICCE TEGONOCIGA CONCUTEGOS | 420 |
| SECOSCOSCO OSCASSITUTO ACCIDENTAN TERACASERE CISCOCCIGA TATTSCORCE | 690 |
| ACTACACGAT ITTUTUGATO GAGGICACTI CGACCITGGA GRACINGTIG CGGAACGCCT | 540 |
| CONTROTTERS OFFERENSES CONTRATORS ASSOCIATED CONTRATOR PROPERTIES | 666 |

84

| EGGRARGEGE ATTEGGRACE ATGGTSTOCK CATCURGETT CTOCKNORGES ITALISTATC CCTSARFOGE GETTTGGED GETCCTTEG ASAATSTGCC TEGGRACE GETCCTTEG TEGGRACECC CTATATGATC GEOGRESTEA TAGOCGACAC CAGGREGAGE CCTACCACAA TEGCGRACEAC CAGCGCTTE TEGGRACET TEGGRACE CAGCGCGAG CCCCCTACACAA TEGCGRACEAC CAGCGCTTE TEGGRACET TEGGRACE CAGCGCGAG CCCCCTTCCC GATATGCGC GEGGGGGGT TASTCCAGGG CTTGGGGTC CGGGGCGAGA CCCCCTTCCC CGGCCCCCGAG GTCGTGGGG TASTCCAGGG CTTGGGGTC CTGGGGTCAC CGCCCCGGGG ACGGCCCCG TCCCTTGGGG TASTCCAGGG CTTGGGGTC CTGGGGTCAC CGCCCCGGGC GGCACCACG CTCCCTTGGGG CACCACGC CTAGGGCGAACACCCCCCT CGCGCCAGGC CCCCACACACGG CACCACCGC CACCACGC CACCACACACCC CGCGCCCGAG CACCACCACCACACACA | | | | | | |
|--|-------------------|--|----------------|-------------|-------------|-------|
| TOCGRACEC STRATGATE GCOCCUTCA TAGOGRACAS CARCEGRAGE COTACCAGAA THE COCATAG CASCECTTO THE COCTEST TORSETAGE CACTEGRAGE GEORGEOODS GATATCEGE GROUNDS THE COCCUTCAT TORSETAGE CACTEGRAGE GCOCCUTCAGE GATATCEGE GROUNDS THE COCCUTATION OF THE COCCUTAGE COCCUTCAGE COCCUCAGE GUESTAGE COMMANDED COTACCAGE COCCUTAGE GCOCCUCAGE ACCOCCUCAGE TOCCUTAGE COMMANDED COTACCAGE ACCUTCAGE GOODS TORSETAGE THE COCCUTAGE GOODS COCCUTCAGE COMMANDED COCCUTAGE GOODS TORSETAGE GOODS THE COCCUTAGE GOODS COCCUTCAGE COCCUTAGE COCCUTAGE TOTACAGE COCCUTAGE GOODS COCCUTCAGE COCCUTAGE COCCUTAGE TOTACAGE COCCUTAGE GOODS COCCUTAGE COCCUTAGE COCCUTAGE (2) INFORMATION FOR SEQ ID BOISS: (A) LENGTH: 13 base pairs (B) TYPE INCLUDED COCCUTCAGE COCCUTCAGE (C) STRATGENSSI SINGLE (D) TOTACCOCCUTAGE COCCUTCAGE COCCUTAGE (XX) SEQUENCE COCCUTAGE ACCUTCATE TOTACCOCCUTAGE (XX) SEQUENCE COCCUTAGE ACCUTCATE TOTACCOCCUTAGE GOODS COCCUTAGE (XX) SEQUENCE COCCUTAGE ACCUTCAGE ACCUTCATE TOTACCOCCUTACCUTAGE GOODS COCCUTAGE GOODS COCCUTAGE COCUTAGE COCCUTAGE COCCU | EGCACAGEGE ATT | RECGAACE ATGETSTC | CA CATCUCOGT | CTCCAGCGCG | TTGAGGTATC | 660 |
| THE CENTURE CASCUSCITE TECCNICS TO TO THE CONTROL CACCIDED GROADDDES 844 GATATOCHE GROUND AND THE CACCIDETY CRECKET CACCIDETY CACCIDATICA CACCIDETY CACCIDETY CACCIDETY CACCIDETY CACCIDETY CACCIDET | CCTSAATCGC GGT | TTTGGCC GGTCCCTC | CG AGAATSTGCC | TGCCGPGTTG | SCTCCGTTGG | 720 |
| GATATOCGC GEOGGCAGC GEOGGCTGT CTCCCGGTC DEGGCCCAAC SCCECTTCGC CGGCCCCAA GTCGTGGGG TASTCCAGG CTTGGGTC CTGGGTCAG GECTGGGGT ACGCCCCCAG GTCGTTGGTG CCGACACUGG CTTGGGGTCA GTGGGGACAG GECTGGGGT TECTCCTAGG GTGGTGGAG GGACCAGCT CTAGGGGACA ACCCCCCCT CGGGTCAGCC GGCACCACCG CGGGGCCGG CACACCCC CTAGGGCAACAGCT GCCGTCAGCC CTCAACACCGA CGGGGCCGGC CGCGGCCCCG ATAATGTTGA AASACTAGGC AACCTTAGGA ACGAAGGAC CGGGGCCGG CGCGGCCCG ATAATGTTGA AASACTAGGC AACCTTAGGA (2) INFORMATION FOR SEQ ID NO:36: (A) LENGTH: 18; base pairs (B) TYPE: BELLET GACACGC CGGGGCCCGG GCGGCCCGCG GCGGGTGTCGG CGGATCGGC GGGGGGTTAA ACGCCAACGC CGGGGCCCGCG GCGGGTGTCGG CGGATCGGC GGGGGGTTAA ACGCCAACGC CGGGGCCCGCG GCGGGTGTCGG CGGATCGGC GGGGGGTTAA ACGCCACCGC GCGGGCCCGCG GCGGGTGTCGG CGGATCGGC GGGGGGTTAA ACGCCACCGC GCGGGCCCGCG GCGGGGCCCC TAACGCTGGT GCCGGCGGAA ACGCCTCGTT CTTAGGGCC GGCGGCCCGCG GCGCGGCCCC TAACGCTGGT GCCGGCGGAA ACGCCTCGTT CTTAGGGCC GGCGAACGCG GCGCTGTCGG CGCACCAATGGT GCCGGCGGAA ACGCCTCGTT CTTAGGGCC GGCGAACGCG 180 GCCINNOCCC TAACGCTGGT GCCGGCGGAACGCG ACGCCTCGTT CTTAGGGCC GGCGAACGCG 180 GCCINNOCCC TAACGCTGGT GCCGCCGGAACGCG ACGCCTCGCT CTTAGGGCC GCCGCCCGCG ACGCCTCGCT CTTAGGGCC GCCGCCCGCG ACGCCTCGCT CTTAGGGCC GCCGCCCGCG ACGCCTCGCT CTTAGGCCAACGCG 180 GCCCCCCCCCCCCCCCCCACCAACGCC ACGCCCCCCCCC | TOCOGACCCC GTA | TATGATE GCCGCCGT | CA TAGOCIJAČAC | CAGCGOGAGG | GCTACCACAA | 780 |
| COGCOCCEAS GTOSTOGGGG TASTCCAGGG CTTGGGGTC GTGGGGTCAGG GGGGGGGGGG | TECCEATCAS CAS | CCGCTTO TGCCGTC6 | CT TOOGGTAGGA | CACCTGCGGC | eecycoccee. | 240 |
| ADGGGCCCG TOCGTTGGTG CCGACACCGG SETTCGCCCA STSGGGCCCGC GSCATTGTGC 100°C THOTCCTAGG GTOGTGGGC GGACCAGCTC CTAGGGCGAC AACCCCCCTC CCGTCAGCC 100°C GGCACCACCG GCGACCAGCTC GTAGGGCGACCAGCT GCCGTCAGCC 114°C CCCAACCCGA CGGGGCGGC GACCAGCCC ATAAFGTTGA AASACTAGGC AACCTTAGGG 112°C ACGAAGGACG GAGGTTTTST GACGATC 122°C (2) INFORMATION FOR SEQ ID NO:36°C (3) SEQUENCE CHARACTERISTICS: (A) LENGTH: 181 base pairs (B) TYRC: DULLELS soid (C) OTRAMEDINESS: mingle (D) TOVOLOGY: 11 hear (Xi) SEQUENCE GGGGGGTGATAA ACGCCACGC GGGGGCCCGC GGGGCCCGC GGGGGCCCGC GGGGCCCGC GGGGGCCCGC GGGGCCCGC GGGGCCCGC GGGGGCCCGC GGGGGCCCGC GGGGGCCCGC GGGGGCCCGC GGGGGCCCGC GGGGCCCGC GGGGGCCCGC GGGGGCCCGC GGGGCCCGC GGGGCCCGC GGGGGCCCGC GGGGGCCCGC GGGGCCCGC GGGGCCCGC GGGGCCCGC GGGGCCCGC GGGGCCCGC GGGGCCCGC GGGGCCCGC GGGGGCCCGC GGGGCCCGC GGGGGCCCGC GGGGCCCGC GGGGCCCGC GGGGCCCGC GGGGCCCGC GGGGCCCGC GGGGGCCCGC GGGGCCCGC GGGGCCCGC GGGGCCCGC GGGGCCCGC GGGGCCCGC GGGGCCCGC GGGGGCCGC GGGGGCCGC GGGGGCCGC GGGGGCCGC GGGGGCCGC GGGGCCCGC GGGGGCCGC GGCGCCGC | CATATOCOGC 000 | POGGCASC GCCGCGTC | ST CESCUSSICS | CGCGCCGAAG | SCOSSTYCEG | 9.00 |
| THOTOCTAGG GYOGTGERIG GRACCAGCTS STARGEGGA AACCCCCCT CEGTCAGCC 1086 GGEACCATCAGGT GABSTUCCTA GGCAGGGCAGCGC GCCGTCAGCT 1140 CTCAACCCCA CGGGGCOGGC CGCGGCCCCC ATAAFGTTGA AASACTAGGC AACCTTAGGR 1286 ACGAAGGACG GAGATITTST GACGATC 1227 (2) INFORMATION FOR SEQ ID NO:36: (A) LENGTH: 181 base pairs (B) TYRE: DULLELS sould (C) STARGEGRACCAGCG GGGGGCCCGC GCCGCCCGCG (A) LENGTH: 181 base pairs (B) TOPOLOGY: 116ear (Xi) SEQUENCE DESCRIPTION: SEQ ID NO:36: GCGGTGTCGG CGGATCGGCC GGGGGCCCGC GCCGCCCGCG GGGCCCGCC TAACGCTGGT GCCGCCGCGA ACCCCTCGTT CTTAGGGGC GGGGGCCCGC GGGGCCCCC TAACGCTGGT GCCGCCGCGA ACCCCTCGTT CTTAGGGGC GGCGGCCCGC GCGCCCCCC TAACGCTGGT GCCGCCGCGA ACCCCTCGTT CTTAGGGGC GGCGGCCCGCG GCGCCCCCC TAACGCTGGT GCCGCCGCGA ACCCCTCGTT CTTAGGGGC GGCGGCCCGCG GCGCCCCCCGC ACCCCTCGT CTTAGGGGC GGCGGCCCGCC (2) INFORMATION FOR SEQ ID NO:37: (A) SEQUENCE CHARACTERISTICS: (A) LENTTH: 296 base pairs (H) TYPS: HECLETC GACG (C) TRANGEDRESS stripte | CGGCGCCGAG GTC | GIGGGG TASTCCAG | GG CTTGGGGTTC | GTGGGATCAC | GOCTOOGGOT | 960 |
| GORACCAGE GENATEGET SMETULETA GREAGETAG GEGARLAGET GEOGTERGET 1146 CTCARACCGA CGGGGGGGC CGCGGGGGCG ATAGGTTGA ARGACTAGGC ARCUTTAGGR 1206 ACCARAGGAC GAGATITTOT GACGATE (2) INFORMATION FOR SEQ ID NO:36: (A) LENGTH: 181 base pairs (B) TYRE: MULLeic seid (C) STRAMEBRONESS: mingle (E) STRAMEBRONESS: mingle (EX) SEQUENCE DESCRIPTION: SEQ ID NO:36: GGGGGGCCGGC CGGGGGCCGGC GGGGGCCGGC GGGGGCCGGC GGGGGCCGCC TARGGGTGGT GCCGGCGGCA AGUCCTGGTT GTTAGGGGC GGGGGCCGGC GGGGCCGGC CACCARTGGT GCCGGCGGCA AGUCCTGGTT GTTAGGGGC GGGGGCCGGC GGGGCCGGC CACCARTGGT GCCGGCGGCA AGUCCTGGTT GTTAGGGGC GGGGGGCCGG GGGGCCGGC CACCARTGGT GCCGGCGGCA AGUCCTGGTT GTTAGGGGC GGGGGCCGGC GGGGCCGGC CACCARTGGT GCCGGCGGCA AGUCCTGGTT GTTAGGGGC GGGGGCCGGC GGGGCCCGCC (A) LENTTYPE INCLIET GAGGCA GGCGGGCCGGC GGCGCGCGCG (A) SEQUENCE CHARACTERISTICS: (A) LENTTYPE; NECLIET GAGG (E) TRANBEDERSS: single | ACGGCGCCCG TCC | CTTGCTG CCGACACC | GG GGTTCGGCGA | OTSGGGGGCCC | 99CATTGT6K | 3026 |
| CTCAACCCCA CGGGGCGGGC CGCGGCGCGC ATAGGTTEA AASACTAGGC AACUTTAGGR 1200 ACCAAGGACG GAGATITTOT GACGATC 1221 (2) INFORMATION FOR SEQ ID NO:36: (3) SEQUENCE CHARACTERISTICS: (A) LENGTH: 18) base pairs (B) TYES: NUCLEIC acid (C) STAMEBRONESS: mingle (D) TOPOLOGY: linear (Xi) SEQUENCE DESCRIPTION: SEQ ID NO:36: GGGCCGGCC CGGGTCGGGC GGGGGCCGGC GGGGCCGGC GGGGGCCGGC GGGGCCGGC CACCAATGGT GCCGCGGGCA ACCCCTGTT CTTAGGGGC GGGGGCCGGC 120 GGGCCGGCC CACCAATGGT GGGGGGCGGA ACCCCTGTT CTTAGGGGC GGGGGCCGGC 120 GGGCCGGCC CACCAATGGT GGGGGGGGG GGTGCGGGGGGGGGG | TECTOCTAGE GTG | GYGGACG GGACCAGC | TO CTAGGGCGAC | ARCCCCCCCT | CCCCTCAGCC | 1000 |
| ACGRAGGACO GROATITOT CACGATO (2) INFORMATION FOR SEQ ID NO:36: (4) SEQUENCE CHARACTERISTICS: (A) LENGTH: 18) base pairs (B) TYPE: INCLEIC socid (C) STAMPBERNESS: single (E) TRANSPORTER SEQ ID NO:36: (XI) SEQUENCE DESCRIPTION: SEQ ID NO:36: GEGETOTICG CEGATOGROE GREGGEGTAR ACCOCACGE GEGEGECEGE GEOGGEGEGEG GEGENOSCOC TAACGETGGT SECCESCEGEA ACRECTOGTT CTTAGGGGC GEOGGEGEGE GEGENOSCOG CACCARTEGT GEGETGGGGC GRITCGTTCAC GRITAGGGCG (Z) INFORMATION FOR SEQ ID NO:37: (A) SEQUENCE CHARACTERISTICS: (A) LENGTH: 28 base pairs (H) TYPE; INCLUSIC gold (E) TRANSPORTESS: single | GGCAGCATOG SCA | ATCAGGT GASCTECC | PA GGCAGGCTAG | GOCAACAGCT | GCOGTCAGCT | 1.140 |
| (2) INFORMATION FOR SEQ IG NO:36: (1) SEQUENCE CHARACTERISTICS: (A) LENTH: 18: base pairs (B) TYES: incleic soid (C) STRANDEDNESS: single (D) TOPOLOSY: linear (XI) SEQUENCE DESCRIPTION: SEQ ID NO:36: GOGGOCOGOCO TANDESTGST OCCOGOGGOG ASSOCIAGOG GOGGOCOGOC GOGGOCOGOC GOGGOCOGOC TANDESTGST OCCOGOGGOG ASSOCIATORY OTTOROGOG GOGGOCOGOC 120 GOGGOCOGOCO TANDESTGST OCCOGOGGOG ASSOCIATORY OTTOROGOGO GOGGOCOGOC 120 GOGGOCOGOCO TANDESTGST OCCOGOGGOCOGOC GOGGOCOGOC GOGGOCOGOC 120 GOGGOCOGOCO TANDESTGST OCCOGOGGOCOGOC GOGGOCOGOC GOGGOCOGOC 120 GOGGOCOGOCO TANDESTGGST OCCOGOGGOCOGOC GOGGOCOGOC GOGGOCOGOC 120 GOGGOCOGOC TANDESTGGST OCCOGOGGOCOGOC GOGGOCOGOC 120 GOGGOCOGOCO TANDESTGGST OCCOGGOCOGOC GOGGOCOGOC 120 GOGGOCOGOCO TANDESTGGST OCCOGOGGOCOGOC GOGGOCOGOC 120 GOGGOCOGOCO TANDESTGGST OCCOGOGGOCOGOC GOGGOCOGOC 120 GOGGOCOGOCO TANDESTGGST OCCOGOCOGOC GOGGOCOGOC 120 GOGGOCOGOCO TANDESTGGST OCCOGOGGOCOGOC GOGGOCOGOC 120 GOGGOCOGOCO TANDESTGGST OCCOGOGGOCOGOC GOGGOCOGOC 120 GOGGOCOGOC TANDESTGGST OCCOGOGGOCOGOC GOGGOCOGOC 120 GOGGOCOGOCOGOC ASSOCIATORY OTTOROGOCOGOC GOGGOCOGOC 120 GOGGOCOGOCOGOC ASSOCIATORY OTTOROGOCOGOCOGOC GOGGOCOGOC 120 GOGGOCOGOCOGOC ASSOCIATORY OTTOROGOCOGOCOGOCOGOC GOGGOCOGOC 120 GOGGOCOGOCOGOCOGOC ASSOCIATORY OTTOROGOCOGOCOGOCOGOCOGOCOGOCOGOCOGOCOGOCOG | CTCAACSCGA CGG | #60006C CGCGGGGG | CG ATAAFGTTGA | AASACTAGGO | AACUTYAGGA | 1200 |
| (i) SEQUENCE CHARACTERISTICS: (A) LENTH: 181 base pairs (B) TYPE: INCLEIC sold (C) STRANDEDNESS: single (D) TOPOLOSY: linear (Xi) SEQUENCE DESCRIPTION: SEQ ID NO:36: QUESTITUTE CONTROL OCCORREGUE ASSOCIATED GEOGRACEGUE GEOGRACEGUE CACCARTEGE SCHOUGESES ATTISTICTAC GEORACEGUE (Z) INFORMATION FOR SEQ ID NO:37: (1) SEQUENCE CHARACTERISTICS: (A) LENTH: 280 base pairs (H) TYPE: INCLUSIC and (E) TRANDEDNESS: single | ACCAMGGACG GAG | NATITIOF GACGATO | | | | 1227 |
| GUGGTGTGGG GEGATOGGG GGTGGTTTAN ACSGCAACGC GGGGGCGGGC GGGGGCGGGG GGACCGGCGC TAALGGTGGT GCCGGCGGCA ACSGCTGGTT GTTGGGGGG GGGGGGGTGGG GGGGGGGGGG CACCAATGGT GGGGTGGGG ACTGCGGGGG ALTTGTCTAN GGCAACGGG 180 (2) INFORMATION FOR SEQ 1D NO:37: (4) SEQUENCE CHARACTERISTICS: (A) LENGTH: 270 base pairs (H) TYPE; INCLUST GOLD (E) STRANDEDESSS; single | (A) (A) (C) | LENGTH: 181 base TYPE: nucleic oc STRANDEDNESS: sin | pal/s | | | |
| GGACCBOCCC TARGETGET GCGGCGGGA AGRICITGET CTTCGGGGC GGGGGGTCCG 120 GGGGNGCCGG CACCAATGOT GGGGTCGGC GGTCGGTGGG ATTGGCTAC GGCAACGGC 180 G 181 (1) INFORMATION FOR SEQ ID NO:37: (1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 290 base paics (B) TYPE; INCLUSE GOLD (C) STRANDEDESSS; single | (xi) SEÇUE | MCE DESCRIPTION: | 38Q 10 NO:36 | < | | |
| GOSENSCOO CACCARGOT GONTTOGGCS SWITCEGUSG ATTESTCIAC GOCAACGGCS 180 (2) INFORMATION FOR SEQ ID NO:37: (1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 290 base paics (B) TYPE; INCLUSIC acid (C) STRANDEDESSS single | ecegrerese esa | ATCOME GOSTOTT | DA ACCOCAACGG | CGGGGCCGGC | ddddccaaca | 60 |
| G (2) INFORMATION FOR SEQ ID NO.37: (1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 290 base paics (B) TYPE; inclusic acad (C) STRANDEDRESS simple | GGACCGGCCC TAX | COSTGGT GCCGGCGGG | DA ACROCTOSTT | OTTORGGGCC | GGCGGGGTCCG | 120 |
| (2) INFORMATION FOR SEQ ID NO.37: (1) SEQUENCE CHARACTERISTICS: (A) LEMOTH: 290 base paics (B) TYPE; inclusic acad (C) STRANDEDRSSS single | GOGGNOCOGG CAC | CARROCT GGROTTGGR | OR SETTOCOGOGO | ATTTGTCTAC | GGCAACGGCG | 180 |
| (1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 290 base pairs (W) TYPE; necleic acad (O) STRANDEDENSS: single | ξÿ | | | | | 181 |
| (A) LEWITH: 290 base pairs (B) TYPE: newhele cold (C) STRAMBEDNESS: strain | (2) INFORMATIO | N FOR SEQ ID NO. ! | 37: | | | |
| | (A) (M) (O) | LENGTH: 298 base TYPE: natieto act STRAMDEDNESS: etr | pairs id | | | |

(xi) SEQUENCE DESCRIPTION: SEQ ID NO.37:

SOSSTSTUGG COGATCOGGC GGGTGGTTGA ACGGCAACGG COGTGTCGGC GGCCGGAGGG

SCHROSCOT CTITESCONT SCHROCOSCO ASSOCIACT CONTENCTA SCHOOLANTS

60

| GCGGCGGCTC CACCGGCGGCA AACGGCGGTC TTGGCGGCGCT GGGCGGTGGC GGAGGCAACG | 180 |
|---|-----|
| OCCODANCE COSCITIONT GOCAACORCO CYAACORTO CONNECCUR RITHUCCOCO | 240 |
| SCAUTCAGAS CECCACOEGO CTOSCHEETE ACGECCOTEA CESECETUAC | 290 |
| (2) INFORMATION FOR SEQ IS NO SEC (1) SECREBED CREARCHESTISTICS: (4) INFORM SEA BOARD PRIES (6) INFES HULLES AND A (6) CHEADEDMESS SINGLE (6) TOPOLOGY: Linear | |
| (xi) SEQUENCE DESCRIPTION: SEQ 10 NO:38: | |
| GATOCAGTEG CATGERGOGT GTCAGTEGAA OCAT | 34 |
| (2) INSORBETION FOR SEC 10 HOLDS: (i) SEQUENCE CHARACTERISTICS: (ii) LENGTH: 155 base pairs (ii) ITPE: nucleic acid (iii) STRANDERIESS: shride (iii) TOPOLOGY: linear | |
| (NI) SEQUENCE DESCRIPTION: SEQ ID NO.39: | 413 |
| CARCACAGCA CALCOCOCOC ALEGARACIA OSCUCIARA COCACOCALIA COCACAGACASOC | 60 |
| TEGOSTOGIC GCCAGCACCC CCGGCACCGC CGACGCCGGA GICGAACAAT GGCATOGICG | 120 |
| TATOCCOACC AFFOCCOCCO GNOCACCOC CAPCO (1) SEQUENCE CHARACTERISTICS: (A) INNOTE: 35 base pairs (B) TTES: rouleic acid (C) STRANDEDMESS: single (D) TOPOLOGY: linear | 355 |
| HRIT REQUENCE DESCRIPTION: SEQ TO NO:48: | |
| ATEGORATICA COGGOCOCCO GGGACCGGGC AGGCCGGNGG GOCCGGGGGG TGG | 53 |
| (2) INFORMATION FOR SEQ ID NO:41: | |
| (1) SEQUENCE CRARACTERISTICS: (A) LEMBTS: 132 been point | |

86

(8) TYPE: nucleic acid

| (C) STAGELEDINGS: SINGLE (D) TOPOLOGY: linear | |
|--|------|
| (xi) SEQUENCE RESCRIPTION: SEQ ID NO:41: | |
| GATCCACCGC GGGTGCAGAC GGTGCCCCCG RCGCCACCCC GACCAGCGGC GGCAAGGGC | 60 |
| GCACCGGCGG CAACGGCGGG AACGCCACCG TOSTCGGNGG GGCGGGGGGG GCCGGGGGCA | 120 |
| AGGGCGGCAA CG | 1.32 |
| (2) INFORMATION FOR SDD ID NO:42: | |
| (i) DEQUENCE CHARACTERISTICS: (A) LENGTH: 132 base pairé (B) TYPE: nucleic acid (C) STEAMDEDMESS: sângle (D) TOPOLOGY: linear | |
| 1%1) SEQUENCE DESCRIPTION; SEQ ID NO:42; GATCHGGGGC CGNAGGGGC GEGGAGGGG REAGGGGGC NAACGGGGC GEGGAAGEER | 66 |
| COMOGGARGA ATECTOGRAS TODACCARTS SOSCOAATOR CSGACROGG: GGCAACUGCS | 120 |
| GRANCSCOG CA | 132 |
| (2) INFORMATION FOR SEC ID NO.43: | 3.32 |
| (i) SECONNE CHARCTERISTICS: (A) LEMETH: TOZ hace pairs (B) TYET: noclaid skid (C) STRANCEIMESS: wingle (D) TOPOLOGY: linear | |
| (#1) SEQUENCE DESCRIPTION: SEQ ID MC:43: | |
| OGSCACGAGG ATOGGTACCO OGOGGCATOG GCAGCTGCOG ATTOGGCGAGC TYTOGCGACC | 60 |
| CHARGARARC CECTACCERA TRECECTECC GRAGIAGESC CATCUSTICS CONTENCES | 130 |
| APGAACOGGO GGCANCAART TAGTOCAGGA ACCTTYCAGT TTAGCGACGA TAATGGCTAT | 380 |
| AGONCTANG DEBOUNDED CARACTERS CONTRIBUTED ACCURATION OF THE CONTRIBUTION OF THE CONTRI | 240 |
| REALTITICAN CARRECCANC DAGGIGGAGE COCCENTESC GUACUTALES ACTENTICO | 300 |
| CONTUNCACO STROUBARDTO ACCOMENTA ARROCCOSC DORACAGERO STRTTETCOS | 360 |

| WO 98/16646 | P("TYT)S07/1826 |
|-------------|-----------------|
| | |

| | 5 | | | | PCT/US97 | /18293 |
|--|---|--|--|-----------------|-------------------------------|--------|
| | | | 87 | | | |
| CCGACARGAS | COSGONATAC | CTG7/20G0/06 | GTGCCAAAGA | SCSGCAGCGT | CTGSCGACCT | 42 |
| SECTORESCAP | CECEGOCIAG | GNOTATESES | AGGITGATUS | GGAGGCTON | ACCOMMONG | 48 |
| ACAACGACGC | CGAAGGAACT | GTGCAGGCAG | ARTCEGGGGGG | GGCCGTCGGA | GGGGACAGTT | 54 |
| CGCCCGAACT | PACCEMENCE | CXXACOCTOG | CLACGGGGGG | TGAACCCAAC | TYCATOGATC | 60 |
| TCAAACAACA | OOCABOGAAG | CTCGAAACGG | GCGACCAAGG | OSCATOROTO | GCGCACTGNG | 66 |
| GCGATGGGTC | GAACACTIRC | ACCOTGACGO | TOCAAGGCGA | CG | | 70 |
| (2) INFORM | NATION FOR S | EQ 10 80:44 | | | | |
| | (A) LBMGTH: (B) TYPE: D: (C) STRANDE: (D) TOPOLOG | onielo acid DNESS: sing | | | | |
| | egoence des | | | | | |
| | PROTETORSS TSCOSTCSSC | | | | | 12 |
| | CTGGTGACAT | | | | | 18 |
| | GTGGCATGGG | | | | | 24 |
| | THE TANGET CHARLES | PLES E MINORINA SE | GGTGGGGGGG | A CONTRACTOR | 888600CCC | |
| | 2009210223212 | MARKET KOOK | AKANMANAAA | Acreanamic. | altra management and a second | |
| AAGTCCAAGG | GTTETCAGCA | | | CCGAGGATUU | TUSTUCCU | 23 |
| AAGTCCAAGG | GTTOTCAGCA ATION FOR SI EQUENCE CHAI (A) LENGTH: (E) TYPE: IN (C) STRANDER (D) TOPOLOS! | CQ ID NO:45: PACTERISTIC: 1058 bace; Poleic sold INESS: sing | s: pairs | CCGAGGASTC | TUSTUCCE | |
| AAGTCGAAGK (2) INFORM (1) S | ATION FOR SI EQUENCE CHAI (A) LEMITH: (B) TYPE: B: (C) STRANGE | eg IO MO:45: PACTERISTIC: 1058 base; Exisic same; Exisic sing! | s s se | | TESTSCES | |
| MAGTCCAAGG (2) INFORM (1) S | ATION FOR SI EQUENCE CHAI (A) LENGTH: (E) TYPE: IN (C) STRANDES (D) TOPOLXS: | EG ID NO:45: SACTERISTIC: 1058 bace; Schein sold UNESS: Singl : limeer | : S: Deite Re ID NOISE | | | |
| AAGTCGAAGG (Z) INFORM (I) S IRI) 5 | ATION FOR SI EQUENCE CHAI (A) LENGTH: (B) TYPE: D (C) STRANDE (D) TOPOLOS: | eg IO NO:45: ACTERISTICS 1058 Date; 201e1c wold NESS: singl F: lineer | schedeere | ; GOACTOCACO | his torsa granc | 23 |
| AAGTCGAAGG (Z) INFORM (I) S IRI) 5 CGGCAGAGG | ATION FOR SE EQUENCE CHAI (A) LENGTH: (E) TYPE: B: (C) STRANGE (D) TOPOLOG: (D) TOPOLOG: EQUENCE DESC ATOGRATICOC | eg ID NO:45: NACTERISTICS 1058 base; Noise base; Noise: single f: linear CRISTION: RE GTUSCOSGGA | Sibelle ACCCCANGE COLORAGE COLORA | ecaccossa | ASTOGAGDAC GSCTCCTACG | 23 |

88

| AGAGCCATGT | GACGGTAGTT | COSTACION | GGSTACTOGG | CGFAFFFCTG | AIGGTCTCGG | 420 |
|------------|-------------|-------------|---------------|-------------|------------|------|
| CGACGTTFAA | CAAGCCCR9C | OCCUPATIONA | coppressions: | ATTGTGGGTT | GROTTGGCTT | 480 |
| TCATCGTGTT | CCAGGGGGTT | GCGGCAGTOC | TOCKOCCTCTT | GGTGGAGAGC | GGCGCTATCA | 546 |
| CCGCGCGCC | 9009096660 | AAGTTOGACC | COTATOGAÇA | GZACGGGGGG | TACGGGCAGT | 600 |
| acgoscasta | CGGGGTEGCAG | CCGGGTGGGT | ACTACGGTCA | CKINGGGTGCT | CAGCAGGCCG | 660 |
| CGGGACTGCA | GTCGCCCGGC | CCGCAGCAGT | CTOCGCAGGC | PCCOGGATAY | GOGTEGCAGT | 729 |
| ACGOCOGCTA | TTOGTCCACT | COGAGOCAAT | CERCOCAGTEG | ATACACTGCT | CAGCCCCCCG | 780 |
| COCAGCCGCC | GGCGCADYCC | GOGTOGCAAC | ANTOGUACCA | GGGCCCATCC | ACGCCACCTA | 846 |
| COURTTICE | SACCTICACC | CCACCACCAC | COSTCASTSC | COGGACCECG | TEGCAGGCTC | 900 |
| GTTCGGCTCC | AGTCAACTAT | TOBARCOCCA | GCGGGGGGGA | GCAGTCGTCG | TOCCCCCGGG | 966 |
| BECCECCEUM | CHAROCOGOC | STICCCCCCGT | CONSTORED | GTGTGCGCGA | AGAGTCANCA | 1020 |
| GGGTGTCASC | AASCGCOGRC | SAYXXIXXIIG | CUGARTIC | | | 1058 |
| | | | | | | |

(2) INFORMATION FOR SEC ID NO:46:

- ii: SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 327 base pairs
 - (5) TYPE: rushaic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ 10 NG:45:

COGGACGAGA GACGATECC GUTAGCCTOS CEGASDASOC AGGIANTITO GARCGGATOT 68
COGGGGACCT GRABACCCAG REGACCAGG TEGASTEGAC GRUBAGTEG TITOCAGGAC
AGTISCOCGO COCGGGGCA AGRCCAGUNI SETECACTE CAAGAACCAG
189
CCAATAAGCA GAACCAGGAA CTCGACGAG TCTCGACGAA TATICCTCAG GCCGGCGTCC 246
RATACTCGAG GCCCGACGAG GAGCAGCAG AGRCCATCT CTCGCAARTG SECTICTCAC
300
CCGCTAATAC GRABAGAAAC GGACCAA 327

(2) INFORMATION FOR DEG IG NG: 47:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 170 bare pairs
 - im) TYPE: nucleic soid
 - (C) STRANDEDNESS: single
 - (D) TOFOLOUY: linear

| (xi) SECUENCE DEDICATPTION: GEO ID NO:47: | |
|--|------|
| CSETCECCAT GRIGGETTE PERACETSA LYGATETTE ACCESSING TEGACATORA | 60 |
| CCHACARCOT GTTGECGTCG GCHARTGTEC CGMACCCGTG SRECTCGGTG ATCTTGTTCE | 129 |
| TOTTCATCAG GAAGIGGACA CCCACCACCO TOCOCTOREN TACCTTTORG | 1,70 |
| (2) INFORMATION FOR SEC ID NO:48: | |
| (i) SEQUENCE CHARACTERISTICS: (A) LORNTH: 13' base pairs (B) FYE: nogles: acid (C) CTRANDERNESS: wingle (D) TOPOLOGY: linear | |
| (wi) SEQUENCE DESCRIPTION: SEQ ID WO:48: | |
| CATTOCOCCG CATGGGGGGT GCCGGCGGCA GCACCGCTGG CSCTGGCGGC AATGGCGGGG | 60 |
| CCGGGGGTEG COCCGGAACC GGTGGGTFGC TCTTCGGCAA CGGCGGTGCC GGCGGGGCACG | 120 |
| Second | 127 |
| (2) INFORMATION FOR DED IN NO:49: | |
| (#) SEQUENCE CRARACTERISTICS: (A) LEMCTH: B1 bese pairs (B) TYPE: canclede acad (C) STRANGEDMESS: atnote (D) TUPCLACY: Linear | |
| (x1) SEQUENCE DESCRIPTION: SEQ 10 No. (6: | |
| COPPOSICIAAG GGCGGCERCCE CCGGCBERCGG GACCGGCGCCG SCCGGCCGCAA ACACCGGCCAA | 60 |
| DESCOUNTED SECURDANCE 5 | 81 |
| (2) INFORMATION FOR SEQ ID NO:50: | |
| (i) SEQUENCE CHARACTERISTICS: (ii) LEMETH 14% bare pairs (iii) TYPES nuclein acid (C) STRANGERNESS single (D) TOPOLOGY, linear | |
| (%) SEQUENCE DESCRIPTION: SEQ ID NO.50: | |

| WO 98/16646 | PCT/US97/18293 | | |
|-------------|----------------|--|--|
| | | | |

| GCAACGGCGG GGCCGCARRY GCCGGCACCAT CCAACCAARRC CGGTAACGGC GGWGCCSGCG | |
|--|---------------------------------|
| | 120 |
| GANACOGYGG TECCOGYGGG CTGATCTGG | 149 |
| (2) INFORMATION FOR SEX FO NO.51: | |
| (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 355 base pairs (B) TYPE: osciolo ecid (C) STRANDERMESS: single (D) TOPOLOGY: linear | |
| (xi) SEQUENCE DESCRIPTION: SEG ID NO:51: | |
| COSCACGAGA TCACACCTAC CGRETEATES MENTCOTESS GRECTESSEE GAGGGTOTES | 66 |
| ACCCCCRANT CCACCCCCTT CTGCCCCCAC CTGCCCACAC CATGCCCCCC CTGCACTGCT | 1.20 |
| TEGRADITACA STCRATICKA GGCCACCIOS TEGROSRACO GGTCGCERAC TIOCROGIEA | 186 |
| CTADGAAAGT CGGCTFCCGC CTOQAGCATT CCTGAACCTT CAAGCCCCGC CCATAACTGA | 240 |
| SETSCATCAT TANGCURETT TECCAGARCA TOUTGROSCH CTCHARACHC SCITCAGGOOD | 380 |
| ACCUTOGOTO OGNOGRADOS CHOCÓTOCAS ASTOCCISCO ACASTOCATO GORGA | 355 |
| (i) SEQUENCE CRARACTERISTICS: | |
| (A) SECTION CONNECTED STITES: (A) LENGTH 99 Eas pairs (B) TYPE: nucleic acid (C) STEANDENESS: cityle [U) TOPOLOGY: licear | |
| (A) LENGTH: 99 take pairs (B) TYPS; unclein and (C) TERMORDERS: single (C) TOPOLOGY: linear (X1) SEQUENCE DESCRIPTION: SEQ 3D NO:52: | 60 |
| (A) LENTH: 99 tase pairs (B) TYPS; uselso and (C) TYPANDEDERS; single (C) TOPOLOGY; ligear (X1) SEQUENCE DESCRIPTION: SEC ED NO:52: ATGUARCACC ATCACCACCA CATECATURG GTURACOCCA ACTITACAUG TOOCRAGGGA | 66 |
| (A) LENTH: 99 tese pairs (B) TYPS: usels: a did (C) TYPANDERS: single (C) TOPOLOGY: ligear (X1) SEQUENCE DESCRIPTION: BED IS NO:52: ATGUATCHE ATCHECATE CATGUATUM STUGARDSCA ACTTURADATE TOPOLANGISA UGATTOSCOG CACTOGCTAT DECESCOGATG SCENACICO DECESTIGAÇÃO CATTOGGSTS | |
| (A) LENGTH: 99 take pairs (B) TYPS: Include and did (C) STRANDHORPS: single (E) TOPOLOGY: ligear (Ki) SEQUENCE DESCRIFTION: BEC ED NO.52: ATGUATURED ATURCATURE CATRUATURE STUGROSCO ACPTRANDED TOPORANGERA CRATTERCOG CACTOGOTAT ORDERCOGATE GRANDHORD DESTRUCTURE COMMUNICATION OF THE | 120 |
| (A) LENTE: 995 tase paid (B) TYPS; inclina add (C) TYPSADEDERS: single (C) TYPSADEDERS: single (C) TOPOLOGY: linear (Xi) SEQUENCE DESCRIPTION: SEQ ED NO.52: ATSCRECAGO ACCAGOCIATO CATGORIOGO ACTTURDADO TOCORAGOGO CGATTOGOCO CACTGOCIATO OCCOGOTATORO COTTOCOGTS COCOGOTOGO COCATOCOGO COCOCOGO TACOCACANO GGCOSPOTOGO CUCCOGTOGO COCATOCOGO GCONCOCOGO COCACOANO GGCOSPOTOGO CUCCOGTOGO COCATOCOGO GCONCOCOGO COCACOANO GGCOSPOTOGO CUCCOGTOGO COCOCOGOGO GCONCOCOGO COCACOANO GGCOSPOTOGO CUCCOGTOGO COCOCOGOGO GCONCOCOGO COCACOANO CUCCOGTOGO COCOCOGOGO GCONCOCOGO | 120 |
| (A) LENTE: 999 tease pair (B) TYPS; inclina oid (C) TYPANDEDERS: single (E) TOPOLOGY: Ligear (X1) SEQUENCE DESCRIPTION: BEQ ED NO.52: ATGUATURGO ATUNCTATURA CATRUATURA STEGRODOGA ACTTURCALS TOGORAGGIA CGATTGGCGG CACTGGCTAT OSCOGGGGTG SCCAGCGCG BOCTGGTGAC CGTTGCGGTG CDGGGGGCCG CCARGGGGG TOGGGGGGCGA GCCCGGGGGACAC GCGCGGCGCG CGGGGGCCG GCGACCCGG GCGACCCGA ACGCGCCACAC CCCACACGGCGCCG CCCCCCCCACACGCCG GCCACCCCG GCCGCACACAC CTCTTGCCCC CCCACACGC | 120 180 240 |
| (A) LENTE: 999 tease pair (B) TYPE; inclein and did (C) TYPENDEDERS: simple (E) TOPOLOGY: Ligear (E) SEQUENCE DESCRIPTION: SEQ ED NO.52: ATSCRICAGO ACCAGGOTAT OCCOGGAGO ACCTTGAGAGO TOGGRAGGGA CGATTGGGGG CACTGGGTAT OCCOGGAGOCA GOCTGGTGAC CGTTGGGGTG CDGGGGGCCG CCARGGDGA TGGGGGGGCA GOCTGGTGAC CGTTGGGGTG CDGGGGGCCG CCARGGDGA TGGGGGGGCA TGGGGGGACAC CTETTGGCCC CDGCGGTCGA CACGGCGAA TGGCGGGACAC CTGTTGGCCC CDCCCGGAGG CACGCCGAACC CACCCGAACC TGTCGGGACC CDCCCGGAGG CACGCCGCA GCCCGAACC CACCCGAACC TGTCGGGACC | 120 180 240 300 |
| (A) EARSTH: 992 kear paid (B) TYPE: nuclein and (C) STEAMDMIRESS: single (D) TOPOLOGY: linear | 120 180 240 200 360 |

| CTTTACGCCA | GCGENCGAAGC | CACCGACTCC | AAGGCCGCGG | cooserress | CPOSSACATS | 600 |
|-------------|-------------|------------|-------------|------------|------------|-----|
| GGTGAGTTCT | ATATOUCCTA | OCCORRACY: | CGGATCAASC | AGGAAACCGT | стосстесяе | 960 |
| GCCAACTGGGG | TGTCTGGAAG | COCUTOGRAT | TACGAAGTCA | AGTTCAGCGA | TOUGAGTANG | 720 |
| CCGAACGCC | AGATCTGGAC | GGGCGTAATC | GGCTCGCCCG | COGCGAACGC | ACCOGACOCC | 780 |
| SGUCCCCCCC | AGCOCTOSTY | TOTOGTATES | CYCGGGACCO | CCAACAACCC | GGTGGACAAG | 840 |
| 00000GGCCA | AGGCCCTGQC | OGAATCGATC | CGOCCTTTGC | POSCOCOGGO | 6008606000 | 900 |
| SCACCGGCTC | CARCMANKA | CCCTCCGGCO | 00000000000 | CCOORGAAGT | CGCTCCTACC | 960 |
| CCGACRACAC | OGACACCGCA | GCGGACCTTA | CCSGCCTGA | | | 959 |
| | | | | | | |

- (2) INFORMATION FOR SECTION NO:531
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 332 amino acida (B) TYPE: amino acid
 - (C) STRANDEDNESS: single

 - (D) TOPOLOGY: linear

THE SEQUENCE DESCRIPTION: BEQ 19 NO:53:

Met His His His His His His Mer His Gln Val Asp Pro Ash Lev Thr Arg Arg Lys Gly Arg Leu Ale Ala Leu Ale Ile Als Ale Met Ale Ser Ale Ber Loo Val Thr Val Als Vei Pro Ale Thr Ale Ann Ale Asp Pro Gin Pro Ala Pro Pro Vol Pro The The Ala Ala Ser Pro Pro Ser The Ala Ala Ala Pro Pro Ala Pro Ale Thr Pro Val Ala Pro Pro Pro 65 Als Ala Ash Thr Fro Ash Ala Glo Fro Gly Asp Fro Ash Ala Ala Pro Pro Pro Ala Asp Pro Asc Ala Pro Pro Pro Pro Val lie Ala Pro Agn Als Pro Cln Pro Vol Arg lik Amp Agn Pro Val Cly Gly Pho Ser Pho Ale Iso Fro Ale Gly Trp Val Glis Ser Asp Ale Ale His Pho Asp Tyr Gly Sex Als Lew Lew Sex Lys Thr Thr Cly Asp Pro Pro the Pro 145 150 165

Gly Gin Pro Pro Pro Tal Ala Asm Amp The Arg lie Vai Lau Gly Arg

Les Asp Gin Lys Les Tyr Ala Ser Ala Gin Ala Thr Asp Ser Lys Ala

Ala Ala Arg Leu Gly Ser Asp Met Gly Gla Phe Tyr Het Pro Tyr Pro

Gly Thr Arg lie Asn Gin Glu Yoz Val Ger Leo Asp Ala Asn Gly Val

Ser Gly Ser Als Ser Tyr Tyr Glu Val Lys The Ser Asp Fro Ser Lys 225 235 240

Pro Aso Cly Gin Ile Trp Thr Sty Val Tie Gly Ser Pro Ale Ale Aso

Als Pro Asp Ala Giy Pro Pro Gin Arg Tep Phe Val Val Tep Lou Gly

Thi Ais Asn Asn Pro Val Rap Lys Gly Ala Ala Lys Ala Lou Ala Glu

Ser lie Ary Pro Leo Vol Bla Pro Pro Pro Ale Pro Ala Pro Ale Pro

Ala Glu Pro Ala Pro Ala Pro Ala Pro Ala Gly Glu Val Ala Pro The

Pro The The Fro The Pro Sis avg The Lew Pro Ala

- (2) INFORMATION FOR SEQ TO MO:54:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 amino action (B) TYPE: amino sold
 - (C) STRANDEDNESS:
 - (D) TOPGLOGY: Linear
 - (%1) SEQUENCE DESURIPTION: SEC ID NO:54:
 - Amp Pro Val Amp Ala Val Ile Am The The Kan Am Tyr Gly Glo Val 1.0
 - Val Ale Ale Lee
- (2) INFORMATION FOR 88Q ID NO:55:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 15 amino acids
 - (R) TYPE: amino soid
 - (C) STRANGEDMESS:
 - (D) TOPOLOGY: Linear

(xi) SECONDER DESCRIPTION: 880 TO NO:55:

Ala Val Glo Ser Gly Met Leu Ala Leu Gly Thr Pro Ala Pro Ser 10

- (2) INFORMATION FOR SEC ID NO:58:
 - (A) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 amino acids
 - (B) TYPE: amino acid (C) STRANDEDNESS:

 - (D) TOPOLOGY: linear
 - (zii SEOCENCE ORSCRIPTION: SEO ID NO:56:

Ala Ala Met Lys Pro Ard The Gly Asp City Pro Lev Glo Ala Ala Lys. 1.0

Gin Gly Arg

- (3) INFORMATION FOR SEQ 1D NO:57:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 15 amino acids
 - (E) TYPE: amino acid

 - (C) STRANDEDMESS: (D) TOPOLOGY: kinear
 - (RI) SEQUENCE DESCRIPTION: SEQ ID NO:57:

Tyr Tyr Trp Cys Pro Gly Sin Pro Phe Asp Pro Ala Trp Gly Pro 10

- (2) INFORMATION FOR SEG 10 NO.58:
 - (4) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 14 amino acids (8) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: Linear
 - (E1) SEQUENCE DESCRIPTION: SEQ 10 NO:58:

Asp fie Gly Ser Glu Ber Thr Glu Asp Gln Glo Kas Ala Val 15

(2) INFORMATION FOR SEQ ID NO:59:

- (1) SEQUENCE CHARACTERISTICS:
 - (A) LEMGTH: 13 amino soids
 - (B) TYPE: amino acid
 - (C) STRANSEDNESS:
 - (D) TOPGLOGY: Linear
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

Als Glu Glu Ser Ile Ser The Saa Glu Xaa Tie Val Pre

- 12) INFORMATION FOR SEC 10 RO: 60:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 17 amino acids
 - (B) TYPE: amino acid IC) STRANDEDMESS:
 - (D) TOPOLOGY: Linear
 - (x1) SEQUENCE SESCRIPTION: SEQ 10 NO: 66:

Amp Pro Clu Pro Ala Pro Pro Val Pro The Ala Ala Ala Ala Pro tro 10

Ala

- (2) IMPORMATION FOR SEC ID NO: 61:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 15 amino acids
 - (A) TYPE: amino sold (C) STRANCEDNESS:
 - (U) TOPOLOGY: linear
 - txi) SEQUENCE DESCRIPTION: SEG ID MO:61:

Ale Pro Lys The Tyr Xsa Glu Glu Leu Lys Gly The Asp The Gly 18 19

- (2) INFORMATION FOR SEQ TO NO:62:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 amino acids (B) TYPE: meine anid
 - (C) STPANDEDNESS:
 - (S) TOPOLOGY: linear

- (wi) SEQUENCE RESCRIPTION: SEQ IS NO:62:
- Asp Pro Ala Sez Ala Pro Asp Val Pro The Ala Ala Sig Gln the Ser 1 16 15

Low Low Asn Asn Lew Als Asp Fig Asp Val Ser Pho Als Asp 20 $$35^{\circ}$$

- (2) INFORMATION FOR SEQ 10 NO:63:
 - (i) SECOENCE CHARACTERISTICS:
 - (A) LEWSTH: 187 amino acids
 - (B) TYPE: Smino acid
 - (C) STRANDEDMESS: single (D) TOFOLOGY: linear

 - (XL) SEQUENCE PROCEPTION: DED IS NO.630
 - The Gly Ser Lep Ass Gin The His Ass Arg Arg Ala Ban Gio Arg tys 1 10 15
 - Asn the the Met Lys Met Val Lys Ser lie Ala Ala Gly Leu The Alæ $20\,$
 - Alo Alo Alo Ile Gly Alo Alo Alo Alo Gly Val The Ser Ile Met Alo 35 ± 6
 - Gly Dly Pro Val Vel Tyr Sin Met Gin Pxo Vel Val Phe Gly Ala Pro 50 60
 - Leu Pro Leu Asp Pro Ala Sar Ala Pro Asp Val Pro Thr Ala Ala Gin 65 75 80
 - Lete The Sez Let let Ash Ser Let Aia Aap Pro Aan Val Ser Phe Ala 85 96 95
 Ash Lys Wly Ser Let Val Clu Siy Siz Ile Siy Siy Yar Siz Ala Arg 100 100 105
 - lie Ala Asp His Lya boo by Lya Ala Ala Gin His Gly Asp Law Pro
 - ion Sor Phe Sex Val The Ass He Gir Pro Alo Ala Ale Gly Sex Ale
 - The Als Aup Val Ser Val Ser Gly Pro Lys Let Ser Ser Pro Val The 145 $\,$ 150 $\,$ 150 $\,$
 - Gln Asn Val Thr Phe Val Asn Glo Cly Sly Trp Met Lou Rer Arp Ale 165 176
 - Sor Ala Met Giu Leu Leu Gin Ala Ala Giy kas 180 186
- (2) INFORMATION FOR SEC 1D NO: 64:

96

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LEMGTH: 148 amino acids
 - (B) TYPE: smins acid (C) STRANGEOMESS: single

 - (D) TOPOLOGY: Linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:64:
- Asp Giu Val Thr Val Giu Thr Thr Ser Val Phe Arg Alo Asp Phe Lou
- Sex Sin Leu Asp Als Pro Ala Gin Ala Gly Thr Glu Ser Ala Ual Ser
- Gly Val Gin Gly Leu Pro Pro Gly Ser Ala Leu Leu Val Val Lys Arg
- Gly Pro Asn Ala Gly Ser Ary Pho Lou Leu Asp Gin Ala 11c Thr Ser
- Ala Gly Arg his Pro hap Ser Arp lie Pho hau Asp Asp Val The Val
- Ser Arg Arg His Ale Glu Phe Arg Leu Glu Aon Asn Glo Phe Asn Vaj
- VAL Amp Val Gly Ser Loo Asn Gly The Tyr Val Asn Arg Glo Pre Val
- Asp Ser Ala Val Leu Ala Ann Gly Asp dlu Val Gln fle Gly Lys Leu 115
- Avg Leu Val Phe Len Tar Gly Pro Lys Glo Gly Glo Asp Asp Gly Sor 135

The Gly Gly Fid

- (2) INFORMATION FOR SEQ TO NO: 65:
 - (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 230 aming scius
 - (B) TYPE: amino sold
 - (C) STRANDEDNESS: single
 - (0) toPGLOGY: linear
 - (Mi) SEQUENCE DESCRIPTION: SEQ 10 NO:65;
 - The Der Adm Ary Pro Als Arg Arg Gly Ary Arg Als Pro Arg Asp The
 - Sly Pro Amp Arg Ger Ala Ser bou Ser bou Val Arg His Arg Arg Gin 20 25

97

Gin Arg Asp Ala Leu Cys hen Ser Ser Thr Gin 11s Ber Ary Gin Ser 35 45 45

Asn len Pro Pro Ala Ala Gly Gly Ala Ala Asn Tyr Ser Arg Arg Asn S0 55

The Asp Val Arg lie Lys lie Pre Met Leo Val The Ala Val Val ber 65 75 80

Let Cys Cys Ser Gly Val Als the Als Als Fro Lys The Tyr Cys Qls $90\,$

Glo bet bys Gly The Asp Thi Gly Gln Als Cys Gln lie Gln Met 867 $105\,$ $116\,$

Amp Pro Ala Tyr Amn Ile Amn Ile Smr Leu Pro Smr Tyr Tyr Pro Amp 113 125

Gln Lys Ser Les Slu Asn Tyr Ile Als Gin Thr Arg Asp Lys the Leu 136 140

Ser Ala Ala Thr Ser Ser Thr Pro Arg Siz Ala Pro Tyr Giu Leu Ash 145 150 160

The Thr Ser Als Thr Tyr Sin Ser Als Ile Pro Pro Ary Sly Thr Gin 165 176 176

Ala Val Val Leu Xas Val Tyr His Aso Ala Gly Sly Thr His Pro Thr 180 195

The Thu Tyr Lys Als the Asp Tru Asp Cin Als Tyr Are Lys Pro 11e 195 200

The Tyr dap Thr Leo Trp Gin Ale Asp Thr Asp Pro Leo Pro Val Val 216 226

Phw Pro lle Val Ala Arg 225 230

- (2) INFORMATION FOR SEQ ID NO:66:
 - (i) SEQUENCE CHARACTERISTICS: (A) LERGIN: 130 amino acids
 - (B) TYPE: amine acid
 - (C) STRANDEDNESS: single
 - (B) TOPOLOGY: linear
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

The Ain Ale Sec Asp Asn Phe Gln Les Sec Sln Sly Sly Sic Ciy Phe I 5

Als the Pro 110 Gly Gln Ala Met Ala Ile Ala Gly Gln 11e Arg Sez 26 26 30

Gly Gly Gty Ser Pro The Val His The Gly Pro The Ale Pho Lee Gly 35

98

Leu Gly Vol Vel Asp Ach Aen Gly Aen Gly Ain Arg Val Gin Arg Val 50 59

Val Gly Ser Alu Pro Ale Ale Ser Less Gly lie Sec The Gly asp Val G5 25

lie The Ale Val Asp Gly Aie Pro lie Amn Ser Ale The Ale Met Aie 85 96

Asp Als Les Ass Gly Bls Als Vro Gly Acr Val Ile Ser Val Ass Trp 100 105

Gin Thr Lyw Ser Gly Gly Thu Arg Thr Gly Ann Val the Lew Ale Glm 115

- Gly Pro Pro Ale
- (2) INFORMATION FOR SEC 15 NO: 67:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 100 amino acids
 - (B) TYPE: amino scid (C) STRANDEDNESS: sincle
 - (D) TOPOLOGY: linear
 - (x1) SEQUENCE DESCRIPTION: SEC ID NO:67:
 - Val Pro Lou Arg Ser Pro Ser Met Ser Fre Ser 198 Cys Leu Ala Ala 1 5 10 15
 - Als Gln Arg Ass Fro Val lie arg Arg Arg Les Ser Ass Fro Pro 20 30
 - Pro Arg Lys Tyr Arg Sex Mor Pro Ser Pro Ala Thr Ala Sex Ala Gly 35 45
 - Met Ala Arg Val Arg Arg Arg Ale lie Trp Arg Gly Ero Ala Thr Xse 50 60
 - Sor Als Gly Met Als Arg Vel Arg Arg Trp Xas Vel Met Pro Xas Val 85 76 76 90
 - lie Glo Ser The Kee Ile Ary Kee Kas Gly Pro Phe Asp Asn Arg Gly 85 95

Ser Glu Arg Lys

- (2) INFORMATION FOR SEQ 10 PO:68:
 - (i) SEQUENCE CHARACTERISTICS:
 - DAI LENGTH: 163 emano acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - 10) TOFOLOXY: Linear

- (wil SEQUENCE DESCRIPTION: SEQ ID NG: 68:
- Met Thr Asp Asp lie Leu Leu lio Asp Thr Asp Glu Arg Val Arg Thr l S 10
- Lam Tor Lev Asn Arg Pro Gin Set Arg Asn Ala Lev Ser Ala Ala Lev $20 \ \ 25 \ \ 30$
- Ile Amp Vai Vai The Leu Thr Oly Ale Amp Pro Val Phe Cya Ala Gly $50\,$
- les Asp les Lys Val Ala Gly Ary Ala Aap Arg Ala Ala Gly Bis Les 65 75 80
- The Als Val Gly Gly his Asp Glo Ald Gly Asp Arg Arg Asp Glo Arg 65 96
- Arg Arg Gly His Arg Arg Ala Arg Thr Gly Ala Val Leu Arg His Pro 105 110
- Asp Ard Leu Ard Ble Ard Pro Leu Ard Ard His Pro ard Pro Gly Gly 125 125
- Ala Ala His Let Sty Thr Sto Cye Vai Let Ala Ala Liyo Gly Arg 135 140
- Sis Arg Rea Gly Fro Val Aep Sis Pro Aep Arg Arg Leu Pro Val Arg 145 150 150
- Asp Arg Arg

(2) INFORMATION FOR SEC 10 NO.89:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 344 amino solds
 - (B) TYPE: amino acid
 - (C) STRANGEDMEDS: single
 - (D) TOPOLOGY: linear
- (xi) SEQUENCE DESCRIPTION: DED IS NO:69:
- Met Lys Phe Val Asm Bis Tim Glu Pro Val Aim Fro Acq Arg Ala Gly
- Gly Ale Vel Ale Dis Vel Tyr Ale Gls Ale Arg Arg Gls Fre Gly Arg $25\,$
- Leu Pro Gia Pro Lou Ala Met Leu Sar Hro Asp Gia Gig Lou Leu Thr 35 45

Als Gly Trp Ale Thr Les Arg Glu The Les Les Val Gly Gln val Pro Arg Giy Arg Lyo Gio Alo Vei Ale Ale Ale Poi Ale Ale Ser Leo Arg 65 75 76 Cys Pro Trp Cys Val Asp Als Bis The The Met Lew Tyr Als Als Giy Glo Thr Asp Thr Ale Ale Ale The Len Ale Gly Tar Ale Pro Ale Ale 105 110 Giy Asp Pro Ash Als Pro Tyr Val Ale Trp Als Als Gly Thr Gly Thr 115 120 Pro Ala Gly Pro Pro Ain Pro the Gly Pro Asp Val Ain Ala Glo Tyr ian Cly Thr Ala Val Gin Phe Bin Phe Sie Als Arp Leo Val Leo Val Leo Leo Asp Glo The Phe Leo Pro Gly Gly Pro Arg Sia Glo Glo Leo 185 Not Arg Arg Ais Oly Gly Leu Val Phe Als Arg Lys Val Arg Ala Glu 180 180 180 His Arg Pro Gly Arg Ser The Arg Arg Leu Glo Pro Arg The Leu Pro Asp Asp Leo Ala Try Ala Thr Fro Ser Glo Pro I) all the Ala Phe Ala Ala leo Ser His His Leo Asp Thr Ala Pro Ala Leo Pro Pro Pro 225 235 240 The Ary Gin Val Vel Arg Arg Val Val Gly See Try Bis Gly Slo Pro 245 250 253 Met Pro Met Ser Ser Acg Trp Thr Ash Glu His Thr Ala Glu Leu Pro Ala Asp Leo Gis Ala Pro Thr Arg Leo Ala Leo Leo The Cly Leo Als 275 285 Pro His Gin Val Thr Asp Asp Asp Val Ala Ala Ala Asg Ber Leu Leu Asp. The Amp Ala Ais Lou Vel Oly Ala Len Ala Try Ala Ala Phe The 305

als Alo Arg Arg He Gly The Trp He Gly Ala Alo Ala Glo Gly Gla

- Val Ser Ang Gin Ame Ero the Gly 349 (2) Impromention for Ang in Mu.78)
 - (i) SECURNOS CHARACTERESTICS:
 (A) LENGTS: 485 amino acids
 (B) TYPE: amino acids

- (C) STRANDEDNESS: Fingle
- DI TOPOLOGY: linear
- (xi) SEQUENCE DESCRIPTION: SEQ 10 NO: 70:

Amp Amp Foo Amp Met Pro Siy Thr Val Ala Lys Ala Val Ala Amp Ala L 10 10

Leu Gly Arg Gly Lie Ala Pro Vei Glu Asp Ila Gin Asp Cya Val Glu 20 25 30

Als Arg bee Gly Glu Als Gly Leu Amp Amp Val Ala Arg Val Tyr iis 40 al Tyr arg Arg Gla Arg Gla Arg Gla Leu Arg Thr Ala Lya Ala Leu Leu 50 50 50 50 50

Arg Glo Arg Tyr Less Less Glo Gla Gly Arg Pro Ala Glu Ser

The City Glo Leu Met Asp Arg Ser Ris Arg Cys Val Als Als Als Glu 100

Asp Gin Tyr Giu Pro Gly Ser Ser Arg Arg Trp Ala Giu Arg Phe Ala 115 125

The law law Arg Arm Lev Clu Fre Law Prv Arm Ser Fre Thi Law Met 130 - 140

Ash Ser Gly The Asp Leu Gly Lea Leu Als Gly Cys Pho Val Leu Pro 145 155 166

The Gin asp Ser Leb Gin Ser lie Phe Ale Thr Lew Gly Gin Ale Air 176 176

Glu Leu Gin Arg Ala Giy Gly Cly Thr Gly Tyr Ala Phe Ser His Leu 185 190

Arg Pro Bla Gly Asp arg Val Aia Ser Thr Gly Gly Thr Ale Ser Gly 195 205

Pro Val Ber Phe Leu Aug Lea Tyr Asp Ser Ala Ala Sly Val Val Ser 210 220

Met Gly Gly Arg Arg Arg Gly Ala Cys Met Ala Val Leo Amp Val Ser 225 230 240

Bis Pro Amp lin Cyx Asp Phe Vel Thr Ala Lys Ala Glu Ser Pro Ser 245 255

Glu Leu Pro Hie Phe Aén Leu Ser Val Gly Val Thr hap Als Fhe Leu 260 270

Arg his Val Glu Arg Asn Gly Lev His Arg Lev Val Asn Pro Arg The 275 280 285

102

- Val Ale 9ro Tar Gly 485 (V) Information for sec to No. 31:
 - (5) SECOENCE CHARACTERISTICS:
 - (A) LEMGTH: 267 amino acade
 - (B) TYPE: amino actd
 - (C) STRANGEDMESS: sincle
 - (D) TOPOLOGY: linear
 - (xi) SEQUENCE DESCRIPTION: SEQ 1D NO:71:
 - Gly Val Lie Val Lee Asp Lau Clu fro Arg Cly Fro Leu fro The Glu 1 18 18
 - The Tys Top Arg Arg Arg Gly Lew Ala Lew Gly Tie Ale Val Val Val 25 30
 - Val Gly 11s Ais Val Ais Nis Val lis Als the Val App Ser Ser Als 35 45

266 26

Ala Pro Pro Pro Glu Ser Pro Ala Gin Gly Gry

- (2) INFORMATION FOR SEQ ID NO:12:
 - (i) SECUENCE CHARACTERISTICS:
 - (A) LENGTH: 97 amino acids
 - (B) TYPE: amine acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (%1) SEQUENCE DESCRIPTION: SEQ 16 NO:72:
 - Leu lle for Thr Gly Lys Ale Ser Rie Ale Ser Leu Gly Val Gir Val
 - Thr Asn Asp bys Asp The Fro Gly Ala bys lie Val Gly Val Val Alg 25 30

104

Gly Gly Als Als As As As Gly Val Fro Lys Gly Val Val Val Thr $35 \,$

Lys Val Asp Asp Asp Pro Ile Ass Ser Als Asp Als Lec Val Als Als 50

Val Arg Ser Lys Ale Pro Gly Ala Thr Val Ala Deu Thr Phe Gln Asp 65 75 PR

Gln

- (2) INFORMATION FOR SEQ ID NG:73:
 - (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 364 amino acido (B) TYPE: amino acid
 - (C) STRANCEDNESS: single
 - (D) TOPOLOGY: linear
 - (81) SEQUENCE DESCRIPTION: SEQ ID NO.73:
 - Gly Ale Ale Val Ser Leu imu Ale Ale Gly Tur Leu Val Leu Thr Ale 1 10 15
 - Cys Gly Gly Gly Thr Ass Ser Ser Ser Ser Gly Ala Sly Sly Thr Ser 20 25 36
 - Gly Ser Val His Cye Gly Gly Lys Lys Glu Len Ris Ser Ser Gly Ser 33
 - The Als Gin Gio Ash Ala Hec Gio Gin Poe Val Tyr Als Tyr Val Arg 50 55 60
 - Ser Cys Fro Gly Tyr Thr Leu Asp Tyr Asn Ala Asn Gly Ser Gly Ala 65 75 98
 - Gly Val Thr Gin Phe Leu Aso Aso Glu Th: Asp Phe Als Gly Ser Asp 85 90 95
 - Val Pro Lau Asa Pro Sar Thr Gly Gln Pro Asp Arg Sar Ala Glu Arg 105 105 110
 - Cys Gly Ser Pro Ale Top Map Lev Pro Thr Val 786 Gly Pro Yie Als 115 120 128
 - lie The Tyr Ash lie bys Gly Yal Sor the Lem Asp Gly Pro $130\,$
 - The The Als Lye Lie the Asn Gly The The The Val Top Asn Asp Pro 14h 150 160
 - Gin lie Gin Ala Leu Ash Ser Sty Thr Asp Leu Pro Pro Thr Pro lie 165 175

105

Ser Vol lie Phe Arg Ser Amp bys Ser Gly The Ser Amp Act the Clo 180

- Lys Tyx Len Asp Giy Vai Ser Asm Gly Ala Tup Gly Lys Gly Ala Sec 195 203
- Giu Thr Phe Sex Gly Gly Val Gly Vel Gly Ala Ser Gly Asa Asa Gly 210 220
- The See Als Leu Leu Gin The The Asp Gly Ser lie The Tyr Asp Glu 225 236 236 235
- Trp Ser Phe Als Vel Gly Lys Gln Leo Ann Met Als Glo Tle Thr 245 255
- Ser Ala Gly Pro Acp Pro Val Ala Lie Th: The Glu Ser Val Gly Lys 265 270
- Why lie Ala Gly Ais Lys II. Not Gly Gim Giy Asn Asp Lou Val Leu 275 280
- Asp The Sec Ser Poe Tyr Ara Dro The Gla Pru Gly Ser Tyr Pro lle 296 300
- Val Lee Ale The Tyr Slo lie Val Cys Ser Lys Tyr Pro Amp Ale Thr 305 310 320
- The Gly the Ala Val Arm Ala Phe Met Gin Ala Ala ile Gly Pro Gly 330 330
- Win Gla Gly Len Asp Gin Tyr Sly Ser lie Pro Leu Pro Lys Ser The 345 956
- Gin Ala kya Loo Ala Ala Ala Vel Aso Blo lie Ser 355 360
- (2) INFORMATION FOR SEQ IS NO. 24:
 - (i) SEQUENCE CRANACTERISTICS:
 (A) LENGTH: 309 amino acids
 - (B) TYPE: amino sold
 - (C) STRANDEONESS; single
 - (D) TOPOLOGY: Linear
 - (xi) SEQUENCE DESCRIPTION: SEQ 10 NO:74:
 - Gin Ala Ala Ala Giy Arg Alo Val Arg Arg the Gly His Ala Glo Arg 1
 - Gin The Hix Gin Asp Ard Los Gis His Gly Cyn Ars Arg Ala Ala Val 26 25 30
 - Val Val Arg Gin asp Arg Ale Ser Val Ser Ale The Ser Ale Arg Fro 35 45
 - Pro Arg Arg His Pro Ala Gin Gly His Arg Arg Arg Val Ala Pro Ser 50 80

Gly Sly Arg Arg Arg Pro His Pro His Nis Vel Slo Pro Asp Asp Arg Arg Asp Arg Pro Ala Lou Leu Asp Arg Thr Sin Pro Ala Giu Bia Pro Asp Dro His Arg Gly Fro Ala Asp Pro Gly Arg Val Arg Gly Arg Gly Arg Leu Arg Arg Val Asp Asp Gly Arg Leu Gin Pro Asp Arg Asp Als Asp Sis Gly Als Pro Wel Arm Gly Arg Gly Pro His Arg Sty Val Gin His Arg Gly Gly Pro Val Phe Val Arg Reg Val Pro Gly Val Arg 145 159 169 Cys Ala Has Arg Arg Gly Haw Arg Arg Val Ale Ale Pro Gly Gle Gly Asp Val ima Ang Ale Sty Lau Ang Vel Clu Ang Lau Ang Pro Val Ala 186 185 185 Als Val Glu Asn Leo Sis Arg Sly Ser Gln Arg Als Asp Cly Arg Vaj 195 200 205 Phe Arg Pro Ile Arg Arg Cly Ale Arg Leu Pro Ale Arg Arg Ser Arg Als Gly Pro Gin Gly Arg Let Ris Let Arp Gly Ale Gly Pro Ser Pro 225 230 Leu Pro Ala Arq Ale Gly Die Gin Glo Pro Ser Ser Ale Gly Gly Arg Arg Ala Gly Giy Ala Glu Arg Ala Asp Pro Gly Gin Arg Giy Arg His 260 265 276 six Oin Dly Gly His Asp Pro Gly Arg Gin Gly Ala Gin Alg Gly Thr Als Gly Vel Als His Als Als Als Gly Pro Arg Arg Als Als Val Arg 295 Asn Ary Pro Ary Ary 305

- (2) INFORMATION FOR SEQ ID NO:75:
 - (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 580 amine acids
 - (B) TYPE: amino soid (C) STRANDEONESS: single

 - (P) TOPOLOGY: Linear

Ser Ala Val Trp Cys Len Aon Gly Phe Thx Gly Arg His Arg His Gly And Cys and Val Ang Ala Ser Gly Trp Ang Ser Sor Asn Ang Trp Cys Ser Thr Thr Alb Asp Cyo Cyo Ale Ser Lys Thr Pro Thr Gin Als Ala 35 Sor Pro Leu Glo Ard Ard The The Cys Cys Sex Pro Ale Wat Gly Cys Arg Pho Arg Ser Pho Pro Val Avg Arg Leu Ale Leu Gly Ale Arg Thr 65 79 90Ser Arg Thr Len Gly Vel Arg Arg Thr Len Ser Gln Tro Asn Len Ser 85 99 Pro Arg Ala Gin Pro Ser Cys Ala Val Thr Tal Glu Set Ris Thr Ris 100 105 Alm Ser Exo Arg Met Ala byo Leu Ala Arg Vai Val Gly ben Val Gin 115 Glu Glu Sin Pro Ser Asp Met Thr Ash Ris Pro Arg Tyr Ser Pro Pro 130 140 Fro Gin Glo Pro Gly Tox Pro Gly Tyr Ale Gin Gly Glo Glo Glo Tax 145 $\,$ Tyr Sex Gln Gln Fbe Asp Txp Arg Tyr Pro Pro Sex Pro Pro Pro Glo 163 Pro The Sin Typ Ang Sin Sec Typ Clu Ale Let Sly Gly The Ang Pro Gly Lou Ile Pro Gly Val Ile Pro The Met The Pro Pro Gly Met 195 200 200 Val Arg Sin Arg Pro Arg Ala Gly Met Leu Ala lle Giy Ala Val Thr 210 215 220 lie Ale Val Val Ser Ala Gly 11e Gly Gly Ala Ala Ala Ser Leo Val 225 236 236 Oly Phe Asn Arg Ala Fro Ala Gly Pro Ser Gly Gly Pro Val Ala Ala Ser Ala Ala Pro Ser lie Pro Ale Ala Aso Met Pro Pro Gly Ser Val 260 265 270 Gio Sin Val Ale Ale Lys Val Val Pro Ser Val Val Per Leo Glo Ter 275 280 285 Amp Let Gly Arg Gin Ser Glu Glu Gly Ser Gly fle lie Let Ser Ala 290 305 Gio Gly Leu lie Leu Thy Aan Aso Mis Val Lie Ala Ala Ala Lys 305 310 310 315 Pro Pro Leu Gly Ser Pro Pro Pro Lys Thr Thr Val Thr Phe Ser Amp

108

3.30 325 Gly Arg Thr Ale Pro Phe Thr Val Val Gly Ala Asp Pro The Ser Asp The Ala Val Arg Val Gin Giy Val Ger Gly Lau Thr Pro Ile Ser Lets Gly Ser Ser Ser Amp Los Arg Val Gly Gin Pro Val Let Ale lle 376 389 Gly Ser Pro Leu Gly Leu Gla Gly Thr Val Thr Thr Gly lie Val Ser Ale Leu Asn Arg Fro Val Ser Thr The Gly Gio Ala Gly Asn Gli Asn The Val bow Asp Als lie Gie The Asp Als Als lie Asm Pro Gly Aso See Gly Gly Ala Leu Val Asn Met Asn Ala Glo Leu Val Gly Val Asn Ser Ala Lie Ala Thr Lee Gly Ala Amp Ser Ala Amp Ala Glo Ser Gly Sor lie Gly Lau Gly Pac Ala lie fro Val Agg Gin Ala Lyc Arg Ile Als Asp Gis Leo lie Ser Thr Gly bys Als Ser Hid Ala Ser Leo Gly Val Gin Val Thr Asn Asp Lye Asp Thr Pro Gly Ale Lye Ile Val Wal Val Val Als Siy Giy Ala Sla Ala Asn Ale Gly Val Pro Lys Gly Val Val Val The Lys Val Asp Asp Asp Pro Ile Asn Sor Ala Asp Ala Leu 530 549 Val Ala his Val Arm Ser byo Ala Pro Gly Ala The Val Ala Leu The 545 556 Phe Gin Asp Pro Ser Gly Gly Ser Arg Thr Vel Gin Vel Thr Leu Gly Lyn Ala Gle Gia

- (2) INFORMATION FOR SEQ 10 NO.76:
 - 141 SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 233 amine acids
 - (B) TYPE: amino acid

 - (C) STRANDEONESS: single (D) TOPGLOSY: linear

(wi) SECOENCE DESCRIPTION: SEC 10 MO:76:

Mot Arm Amp Gly Lys Ard Ala Val The Sor Ala Val Lee Val Val Lee 1 5 10

Gly Ala Cys Las Als Lou Trp Lou Ser Gly Cys Ser Sor Pro Lys Pro

Amp Ala Glu Glu Gin Gly Val Pro Val Ser Pro Thr Ala Ser Amp Pro 35 40 45

Aks Leu Leu Aks Glu lie Ary Gla Ser Leu Asp Aix Thr Lys Gly Leu 50 66

Thi Ser Val His Val Ala Val Arg Thr Thi Gly Lys Val Asp Ser Leu 65 76 40

Let Gly Ile Thr Ser Ala Asp Val Asp Val Arg Ala Ash Pro Let Ala 85 $90\,$

Als Lyn Gly Val Lyn Thr Tyr Ann Amp Glu Gln Gly Val Pro Pan Ary 190 110

Val Gin Gly Asp Asn lie Ser Val Lye hou Phe Asp Asp Trp Ser Asn 125 128

Leu Gly Ser Ile Ser Giu Leu Ser Thr Ser Are Val Leu Asp Fro Ala 130 140

Also Oly Val Thr Cla Leu Leu Ser Gly Val Thr Asa Leu Gin Als Gin 145 150 155 168

Gly The Olu Val lie Aap Gly lie Ser The The Lys lie The Gly The 165 175 175

Ile Pro Bla Ser Ser Val Lys Met Lou Asp Pro Gly Ala Lys Ser Ala 189 195 Rrg Pro Als Thr Val Trp Ile Ala Gln Asp Gly Ser Rio His Leu Val

Arg Ale Ser lie asp bed Gly Ser Gly Ser lie Glo Lou Thư Gin Ger

Lyd Trp Ass Gis Pro Vai Ass Val Asp

- (2) IMPORMATION NOW DEG TO BOLTZI
 - (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 66 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDMESS: single (D) TOPOLOGY: ligger

 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO.77:
 - Val lie Asp lie lis Gly Thr Ser Pro Thr Ser Trp Gto Gin Ala Ala

110

\$ 30

Als Gin Als Val Gin Aro Als ary Asp Ser Val Asp Asp 11e ary Val 20 - 30

Als Ary Val IIo Glu Gln Asp Set Ala Val Asp Ser Ala Sly Lye 110 35

The Tyr Arg Ile Lys Leu Glu Val Ser The Lys Met Arg Pro Als Gin 50 55

Pro Arg

- 12) IMPORMATION FOR SEQ ID NO.78:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LBNGTH: 69 amino solds
 - (B) TYPE: smino acid (C) STRANDEDMESS: single
 - (B) TOPOLOGY: linear
 - (xi) SECOUNCE DESCRIPTION: SEC 10 NO:78:
 - Val Pro Pro Ala Pro Pro Leu Pro Pro Leu Pro Pro Ser Pro Ile Bex 1 10 15 15
 - Cys Als Ser Pro Pro Ser Pro Pro Les Pro Pro Ala Pro Pro Val Ala 26 29 30

Pro Gly Pro Peo Met Pro Pro Lea Asp Pro Trp Bro Pro Ala Bro Pro 33

Let Pro Tyr Ser The Pro Ero Gly Ala Pro Lett Pro Pro Ser Pro Pro 50

Ser Pro Pro Leu Pro

- (2) IRRORMATION FOR SEC 10 NO:79:
 - (1) SEQUENCE CHARACTERISTICS:
 - (A) LEWGTH: 355 amino acids
 - (8) TYPE: amino acid
 - (C) STRANGEDWESS aimole
 - (D) TOPOLOGY: linear
 - (xi) SEQUENCE DESCRIPTION: SEQ 15 NO: 19:

Mot Ger Amb Ser Ary Arg Arg Ser Lea Arg Trp Ser Trp Lea Lea Ear 1 10 15

Val Leu Ala Ase Val Cly Leu Gly Leu Ala Tur Ala Fru Ala Gin Ala

Als Pro Pro Als ben Ser Gin Asp Arg the Als Asp the Pro Ale ben Pro Seu App Pro Ser Ala Met Val Ala Gin Val Ala Pro Gia Val Val 50 99 86 Asn lie Asn Thr Lys Lea Gly Ty: Asn Asn Als Val Gly Als Gly Thr 65 75 76 GLy lie Val lie Asp Pro Asn Gly Val Val Lou The Asn Ash His Val The Ala Gly Ala The Asp The Asc Ala Phe Ser Val Gly Ser Gly Glo 105 105 The Tyr Gly Val Amp Val Val Gly Tyr Amp Arg Thr Gin Amp Val Ala Val Leu Gin Lea Ard Cly Sie Gly Gly Leu Pro Ser Ale Ale Ile Gly Gly Gly Val Ala Val Gly Glo Pro Val Val Ala Met Gly Aco Ser Gly 145 156 Gly Glo Gly 91y Thr kro Arg Ala Vai Pro Gly Arg Val Val hla Leu 165 170 175 Gly Glo Thr Val Glo Als Ser Asp Ser Leo Thr Gly Ala Glo Glo Thr 186 186 Let Apr Cly Let Ile Sin Phe Asp Als Als lie Nis Pro Cly Asp Ser 195 205 Gly Gly Fro Val Val Asn Gly Leu Gly Glo Val Val Gly Met Asn Thr 710 225 Ala Ala Ser Asp Asn Phe Gli Leo Ser Gin Gly Gly Glo Gly the Ala 225 230 230 ils Pro lie Gly Gln Ala Met Ala The Als Gly Gln the Arg Ser Gly 255 Gly Gly Ser Pro Thr Val His lie Gly Pro Tor Ala Phe Leo Gly Leo 260 270 Gly Val Val Asp Asn Asn Gly Asn Gly Ala Ash Val Gla Arg Val Val 275 280 285 Gly Ser Ala Pro Ala Ala Ser Lou Gly Ile Ser Thr Gly Asp Val The 290 906 The Ale Val Asp Giv Ale Pro lie Asn Sec Ale The Ale Mer Ale Asp Als Lou Asn Gly His His Pro Gly Asp Val Ile Ser Val Asn Trp Gle 325 The Lya Sor Gly Gly The Arg The Gly Arm Val The Leu Ale Glu Gly 340 Pro Fro Ala

112

355

(2) INFORMATION FOR SEC TO MO:00:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 205 amino acide

(B) TYPE: amino soid

(C) STRANDEDNESS: single

(D) TOPOLOGY: Linear

Ixil SEQUENCE SESCRIPTION: SEQ ID NO:85:

Ser Pro ige fro Asp Ale Glu Glu Gle Gly VAL Pro Val Ser Pro Thr 1 10 15

Alo Svr Asp Pro Ala heb leb Als Glo Ile Arg Glo Spr Leb Asp Ala 20 25 25 30

The Lye Gly Lee Thr Ser Val Ris Vol Als Val Arg The The Gly Lye

Val Asp Ser Leo Leo Ciy He The Ser Als Asp Val Asp Val Arg Ain 50

And Pro Leo Ala Ala Lyo Gly Val Cye The Tor Asn Asp Glo Gln Sly 65 75 80

Vol Pro Pho Arg Val Gin Gly App Asn Tie Ser Val Lys Leu She Asp 90 95

Asp Trp Ser Asc Lou Gly Ser Lio Ser Gin Leu Ser Thr Ser Arg Val 160 105 110

Lou Asp Pro Ala Ala Gly Val Thr Glo Lou Lou Eex Gly Val Thr Asm 125 126

Let Gin Ala Gin Giy Thx Giu Val ile Axp Gly lie Ser The the Lys $130\,$

Ile Thr Sly Thr fie Pro Ale Ser Ser Val Lys Met Leu Amp Pro Gly 145 155

Als Lys Ser Als Arg Pro Als Thr Vel Trp fle Als Oln Asp Sig Ser 165 170 175

Mis His Leu Val Arg Ale Ber lie Acp Leu Gly Ber Gly Sor lie Gi: 185 190

Let The Gle Ser Lys Tip Ase Glu Pro Val Ase Val Ase 195 200 205

(2) INFORMSTION FOR SEQ 10 NO.81:

- (L) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 285 amino acids
 - (B) TYPE: amino wold
 - (U) STRANDEONESS: Single
 - (0) TOPORCESY: Linear

- (MI) SEQUENCE OBSCRIPTION: SEQ ID NO: 81;
- Gly Asp Sec Phe Trp Ala Ala Asp Sin Sec Ala Arg Gly Phe Val
- Let Giy Als Thr Als Gly Arg Thr Thr Let Thr Gly Glt Gly Let Gln 25 30
- Sis Ala Asp Giy Sie Ser Lee Lee Lee Asp Ala Thr Asp Pro Ala Val 35 45
- Val Ala Tyr Asp Pro Ala Phe Ala Tyr Glu 11e Gly Tyr fie Xau Glu 50 50
- Ser Gly Leu Ale and Met Cys Gly Ulu Aso fro Glu Aso fle rhe Pho 65 70 75 80
- Tyr lie The Val Tyr Asc Glu Pro Tyr Val Gle Pro Pro Glu Pro Glu 85 90
- Asm the Amp Pro Glu Gly Val Leu Gly Gly lie Tyr Arg Tyr Bis Ala 100 105
- Als The Glu Gin Arg The Asm bys Ess Gin the Leu Als Ser Gly Val
- Als Met fro Ala Ala Leu Arg Ala Al δ Gln Met Leu Ala Ala Glo Trp 130 \$130
- Asp Val Als Als Asp Val Trp Ser Val The Ser Trp Gly Glu Len Asp 143 150 150
- Arg Asp Gly Val Val Ile Glu Thr Glu Lye Leu Arg His PiA Asp Arg 165 $$170\ \,$
- Pro Ala Gly Val Pro Tyr Vol Thr Sag Ala Leu Gla Asn Ala Arq Gly 185 190
- Fro Val 11s Ala Val Sex Asp Trp Met Arg Ala Val Pro Glo Glo Ille 195 200 205 205 205 Arg Pro Trp Val Pro Gly The Tyr Leo The Leo Gly The Exp Gly Phe 210 226
- Giy Phe Ser Asp Thr Arg Pro Ald Gly Arg Arg Tyr Phe Aon Thr Asp 225 235 240
- Ala Glu Sex Slu Vel Gly Arg Gly Phe Gly Arg Gly Trp Pro Gly Arg 245 250 255
- Arg Val Arm lie Asp Ero Phe Gly Ais Gly Arg Gly Pro Pro Als Gla 265
- Leu Pro Gly Phe Asp Glu Gly Gly Gly Leu Arg Pro X3a Lys 275 285
- (3) INFORMATION FOR SEG 1D MO:82:

- (a) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 173 amino ocida
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- 185) SEQUENCE DESCRIPTION: SEQ ID NO:82;
- Thr Lys Phe His Als Leu Met Glu Glu Gin Ile His Asm Glu Phe Thr I 10 15
- Ala Ala Gin Gin Tyr Val Ala 11e Ala Val Tyr Phe Asp Ser Glu Asp 25 30
- Leu Pro Gla Leu Ala Lys His Phe Tyr Sex Gia Ala Val Glu Glu Arg 35 45 45
- han His Ale Met Het Leu Val Gin His Leu Leu Asp Arg hap Leu Arg 50 60
- Val Glu lie Pro Gly Val Amp Tor Val Ang Amp Gin Phe Amp Ang Pro 65 70 76 80
- Ary Glu Ala Leu Ala Léu Ale Leu Amp Glu Glu Arg Ykr Val Thr Amp 95 95
- Glo Val Gly Arg law The Als Val Ala Arg asp Glo Gly asp the Lew 105 105
- Gây Cân Gân Phe Het Gân Try Phé Lou Gân Gân Gân Tân Gân Gân 981 125 125
- Ale Let Mer ale The Lot Val Are Val Ale Asp Arg Ale Sly ale Ase 130
- lem Phe Glu Leu Glu Aso Phe Val Ale Arg Glu Val Asp Val Ale Pro 145 150 150
- Ala Ala Sar Giy Aik Pro Bis Ala Ala Giy Giy Arg Leb 165 170
- (2) INFORMATION FOR SEQ 10 MO(83:
 - (1) SEQUENCE CHAPACTERISTICS:
 - (A) LENGTH: 107 amino ecids
 - (B) TYPE: saino acid
 - (C) STRANDEDNESS; single
 - (D) TOPOLOGY: linear
 - (ki) SEQUENCE DESCRIPTION: SEC ID NO:83:
 - Arg his Asp Siu Arg Lys Asm Thr Thr Met Lys Met Val Lys Jer 11e

PCT/US97/18293

Als Ain Gly bed for Als Als Als Als Gly Als Als Als Als Bly

Val Thr Ser Ile Met Ala Gly Gly Pro Val Val Tyr Gla Met Gla Pro

Vol Val Phe Siy Als Pro Lee Pro Lee Asp Pro Xee Ser Ala Pro Xas

Val Pro Thr Ale Ale Gio Tep Thr Kas Leo Leo Ass Kee Leo Xaa Asp

Pro Asn Val Ser Phe Xee Asn Lys Gly Ser Lea Val Glu Gly Gly Ile

Gly Gly Mas Glu Gly Xes Mas Arg Arg Mas Glo

- (2) INFORMATION FOR SEQ ID NO:04:
 - (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 125 amino acids (B) TYPE: amino soid
 - [C] STRANDEDNESS: single (D) TOPOLOGY: Linear
 - (xi) SECREMOE DESCRIPTION: SEC ID NO:84:
 - Val Len Ser Val Pro Val Gly Asp Gly Phe Trp Xas Arg Val Val Ass
 - Pro Leu Gly Gln Pro Ile Asp Gly Asp Gly Asp Val Asp Ses Asp The
 - Arg Arg Ale Leu Glu ben Gin Ale Pro Ser Val Vai Kas Arg Gin Gly
 - Val Lys Glu Pro Lec Xna The Gly lis Lys Ala lie Asp Ala Net The
 - Fro Its Cly Arg Gly Sin Arg Sin Leu Its Tie Sly Asp Arg Lys Thr
 - Gly Lys Asn Arm Azq Leu Cys Arg The Pro Ser Ser Am Sin Arm Gla
 - Glis ben Gly Val Arc Trp Tle Fro Arg Ser Arg Cys Ala Cys Val Tyr 100 105
 - Vol Gly His Are Ale Arg Arg Gly The Tyr His Arg Arg
- (2) INFORMATION FOR SEQ ID NO:85:
 - (5) SEQUENCE CHARACTERISTICS: (A) LEEGTE: 117 smins acids
 - (B) TYPE: amino acid
 - (C) STRANSHIMESS: single

(D) TOPOLOGY: lines:

- (xi) SEQUENCE DESCRIPTION: SEO ID NO:85)
- Cyn Asp Als Yai Mar Gly Phe Lou Gly Diy Als Gly Ero Leu Als War l
- Val Amp Gin Glo Leu val for Arg Val Fre Gin Gly Txp Ser Phm Ala 29
- Gin Ale Ale Val Pro Vel Val The Lou Thr Ale Trp Tyr Gly Lew 35
- Ala Aan Len Ala Glu Ile Nys Ala Gly Glu Ser Val Len 11e 91s Ala 50 58
- Gly Thr Gly Giy Vel Gly Net Ale Ais Val Gin led Ale Arg Gln Trp 65 70 75 80
- Gly Val Glo Val Phe Val Thr Ala Ser Arg Gly Lyo Trp Asp Thr Less 85 90 95
- Ard Ala Xoa kas the Asp Asp Xon Pro Tyr Ard kas Mye Pro His Xoo 100 110

Arg Ser Ser Xea Gly 115

- (F) INFORMATION FOR SEQ 30 NO:86:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 103 amine acids
 - (8) TYPE: spine soid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: Linear
 - (xi) SROGENCE SESCRIPTION: SEO ID NO.88:
 - Mat Tyr Arg the Ala Cys Arg Thr Leu Met Leu Ala Ria Cys IIn Leu I 16 15
 - Ala Thr Gly Val Ala Gly Lou Gly Val Gly Ala Gln Ser Ala Ala Gln 20 25 30
 - The Ala Fro Val Pro Amp Tyr Tyr Try Cys Pro Gly Glo Fro the Amp 35
 - Pro Ala Trp Rly Fro Asn Trp Asp Pro Tyr Thr Cys his Asp Asp Phe 50
 - His Arg Amp Net Amp Cly Pro Amp His Ser Arg Amp Tyr Pro Cly Pro 55 75 86
 - lle Lou Glu Gly Pro Val Lou Asp Asp Pro Gly Ala Ala Pro Pro Pro

117

85 90 95

Fro Ale Ale Gly Sly Gly Ale

- (2) INFORMATION FOR BEQ 10 NO:87;
 - (i) SEQUENCE CHARACTERISTICS:
 (A) LEMGTH: 88 amino acide
 (B) TYPE: smino acide
 (C) STRANGEDNESS: single
 (U) TOPCLOGY: ligner
 - (xi) SEQUENCE DESCRIPTION: SEO ID NO: 87:

Vol Gin Cys Arg Vol Typ Lea Giu Tie Gin Typ Arg Gly Mot Leu Gly 15

Ala Asp Olm Ala Arg Ala Gly Gly Fro Ala Arg fle Trp Arg Clu His 20 28 30

Ser Met Ala Ala Met Lys Pro Arg Thr Gly Amp Gly Pro Les Siu Ala 35 46 45

The Lym Glo Gly Arg Gly lie val Met Arg tal Pro Leo Olu Gly Gly SC 55

Giy Arg Leu Val Vai Gia Leu The Pro Asp Gio Ala Ala Ala Lau Giy 65 70 75 80

Amp Glu Len Lys Gly Vel The Son 86

- (2) INFORMATION FOR SEQ ID NO:88:
 - (i) SEQUENCE CHARACTERISTICS:
 (AV LEMBIS: 95 amino acide
 (B) TIPE: amino acide
 (C) STRANGEDMESS: single
 (D) TOPOLOGY: limear
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:00:

The Asp Ala als the Leo als Gin Glu als Giy Asn the Giu arg field 1

Ser Gity Amp Les Lys The Cin Tie Asp Gin Val Gid Sor The Ala Gly 25 30

See Lee Gle Gly Gle Trp Arg Gly Ale Ale Gly Thr Ale Ale Gle Ale 35 45

Ala Val Val Arg Phe Gin Giu Ala Aia Asn Lys Gin Lys Gin Glo Leu 50 60 Asp Glo Ile Ser Thr Aso Ile Acg Glo Ale Gly Val Glo Tyr Ser Arg

Als Asp Glo Gho Glo Gin Glo Als Leo Sex Ser Glo Het Gly Phe

- (2) INPORMATION FOR SEQ IS NO:89:
 - (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 186 amino soids
 - (B) TYPE: emine soid
 - (C) STRAMDEDNESS: single
 - (3) TOPOLOGY: linear
 - (xi) SECUENCE DESCRIPTION: SEC 18 NO:80:

Met The Gin Sex Sin The Val The Val Asp Gin Gin Gin I e Leu Ann

Arg Ala Asn Glu Val Glu Ala Pro Met Ala Asp Pro Pro The Asp Val

Pro The Thr Pro Cys Git Let The Kas Xas Lys Asn Als Ale Gin Gin

Kas Val Leu Ser Als Asp Asm Met bry Gto Tyr Leu Als Als Cly Als

Lys Giv Arg Gio Arg Leo Ala Thr Ser Leo Arg Aso Ala Ala Lyo Xaa

Tyr Gly Glu Val App Glu Glu Als Als The Als two Asp Ass Asp Qly

Gin Gly Thr Val Gin Als Glo Ser Als Gly Als Val Cly Gly Asp Ser 100 105 115 Ser Ale Gin Les Thr Asp The Fro Arq Val Ale Thr Ale Gly Gio Pro

Aso Poe Bet Asp Leo Lye Slu Ala Ala Arg Lys Leo Glu Thr Gly Asp

Glo Giy Ala Ser Leu Ala His has Gly Asp Gly Trp ass The Ras Thr

155

Lou Thr Lou Gin Gly Asp

- 121 INFORMATION WOR SEC TO BO: 95:
 - (i) SECURNCE CHARACTERISTICS:
 - (A) LENGTR: 5 amino eqids (B) TYPE: amino soid

 - (C) STRANSEDSESS: sibule
 - (D) TOPOLOGY: Linear

INI) SEQUENCE DESCRIPTION: SEQ ID NG: 90:

Arg Ala Giu Arg Het.

- (2) INFORMATION FOR SEQ ID NG:91:
 - (i) SEQUENCE CHARACTERISTICS: (A) LEWSTR: 269 amino acids
 - (8) TYPE: amino sold (C) STRANGEDMESS: aingle
 - (D) TOPOLOGY: linear
 - (EI) SEQUENCE DESCRIPTION: SEQ IS NO.91:
 - Val Ala Trp Met Sur Vol The Ala Gly Glu Ala Glu Leu Inz Ala Ala

119

PCT/US97/18293

- Glo Val Arg Val Ala Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu Thr
- Val Pro Pro Pro Val Tie Ale Glu Aso Aug Ale Glu Léu Met ille Leu 35 40 40
- The Ala Thy Aen Loss best Gly Glo Asn Thr Pro Ala Lie Ala Val Awn 56
- Giu Ala Gin Tyr Gly Gin Net Trp Ala Gin Asp Ala Ala Het Phe
- Gly Tyr Ale Ale Ale Thr Ale Thr Ale Thr Ale The Len Leu Pro Phe
- Glo Glo Als Fro Glo Mer The Ser Als Gly Gly Leo Leo Glo Glo Als
- Bla Bla Val Glo Glo Bla Ser Asp Thr Als Als Ala Asa Glo Leu Met
- Ash Ash Val Pro Gin Ala Lau Lys Gin Leu Ale Gin Pro Thr Gin Sty
- The The Pro See See Lys Lau Gly Gly Lau Tre Lys for Val See Pro 146 150 166
- His Arg Sar Pro lie Sar Aon Met Val Scr Met Alo Asr Ass His Met 185 170 179
- Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Ley Ser Ser Mer 185 185
- Lou Lyo Gly Phe Ala Pro Ala Ala Ala Ala Cin Ala Vai Gin The Ala 185 205

120

Ala Gin Ain Giy Vai Ary Ala Met Ger Ser Leu Giy Ser Ser ben Giy 210 225

Ser Ser Gly Leu Gly Gly Gly val ale Ala Asn Leu Gly Arg Ala Ala 225 280 280

Ser Val Ard Tyr Gly Sie Ard Asp Cly Gly Lys Tyr Ala Kaa Ser Gly 245 255

Arg Arg Asn Gly Gly Pro Ala

(2) INFORMATION FOR SEC ID NO. 92:

- (3) ESQUENCE CHARACTERISTICS:
 - (A) LEMGTH: 303 amino acids
 - (B) TYPE: smine acid (C) STRANGEDNESS: single
 - (D) TOFOLOGY: Linear
- (81) SEQUENCE DESCRIPTION: SEQ TO NO: 92:

Met The Tyr Ser Pro Gly Asn Pro Gly Tyr Pro Gle Ala Gle Pro Ala I 10 15

Gly Ser Tyr Gly Gly Val Thr Pro Ser Phe Als His Als Asp Gip $26 \ \ 26$

Als Sex Lys Leu Pro Met Tyr Leu Asn lie Als Val Als Val Leu Gly 45

Let Ala Ala Tyz Phe Ala Sor Phe Gly Fro Met Phe The Let Ser The 50 55

Gin Leu Gly Gly Sly Asp Sly Ais Val Ser Gly Asp Thr Gly Leu Pro $90\,$

Val Gly Val Ala Leu Leo Ala Ala Leo Ala Gly Val Val Leo Val 85 95

Pro Lys Ala Lys Ser Him Val Thr Val Val Ala Val Leo Gly Val Leo 105 116

Gly Val the Low Mot Val Sor Ala Ter Fhe Asn Lys Pro Ser Ale Tyr 115 126 126

Ser The Gly Trp Ala Leu Trp Val Val Leu Ala Che ile Val Phe Gle 130 140

Als Val Ala Ala Val Leu Ala Leu iou Val Gio Thr Gly Alo Ilo Thr 145 155 165

Ala STO Ala Pro Arg Pro Lys Phe Asp Pro Tyr Wly Cin Tyr Cly Arg 185

Tyr Gly Gln Tyr Gly Gln Tyr Bly Val Glo Pro Gly Gly Tyr Tyr Gly 185

121

Gin Gin Giy Ala Gin Gin Ala Ala Giy Len Gin Box Pro Giy Pro Gin 195 205

Gln Ser Pro Gin Pro Pro Giy Tyr Gly Ser Gin Tyr Gly Gly Tyr Ser 218 220

Ser Ser Pro Ser Gin Ser Gly Ser Gly Tyr Thr Ala Gin Pro Pro Ala 229 235 240

Clin Pro Pro Ala Gle Ser Gly Ser Glo Glo Ser His Gin Gly Pro Ser 245 255 255

The Pro Pro The Gly Phe Pro See Phe Ser Pro Pro Pro Pro val Ser 360 205 270

Als Gly Thr Gly Ser Gin Als Gly Ser Als Pro Val Asn Tyr Ser Asn 275 280

Pro Ser Cly Gly Clu Gin Ser Ser Ser Pro Cly Cly Ala Pro Val 295 300

- (2) INFORMATION FOR SEQ TO NO: 93:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 amino acids
 - (5) TYPE: amino acid (C) STRANDEDOWSG: single
 - (D) TOPULOGY: Linear
 - (x)) SEQUENCE LESCRIPTION: SEQ ID NO: 93:
 - Gly Cys Gly Glo The Asp Ala Ala The Lem Ala Glo Glo Ala Gly Asp I $_{\rm 5}$

Phe Glu Arg lie Sex Sky Asp Leu Lys Thr Gin Ile 20 25

- 121 INFORMATION FOR SEC 10 NO:94:
 - (4) SEQUENCE CHARACTERISTICS: (A) LENGTH: 16 amino soigs
 - (B) TYPE: amino sold
 - (C) STRANDEDNESS: single
 - in Torology: linear
 - (#1) SEQUENCE DESCRIPTION: SEQ 10 NO:94;

Asp Gin Val Glu Ser Thr Ala Gly Ser Lau Gln Gly Gin Trp arg Gly i $$5\,$

- (2) DEFORMATION FOR SEQ ID NO. 95:
 - (i) SECTEMBE CHARACTERISTICS:

- (A) LENGTH: 27 amino acids
- (B) TYPE: aming acid
- (E) SYRANDERMESS: single
- (D) TOPOLOGY: linear
- (NI) SEQUENCE DESCRIPTION: 35Q ID 80:95;

Sly Cys Gly Ser Inx Als Gly Ser Leu Gin Sly Gin Frp Arg Gly Ata

Ala wiy Thi Ala Ala Cin Ala Ala Val Val Arq

- (2) INFORMATION FOR SEG ID NO. 96:
 - (i) SEQUENCE CEARACTERISTICS:
 - (A) LENGIB: 27 amino ocida (B) TYPE: emino acid
 - (C) STRAMMEDNESS; single
 - (D) TOPOLOGY: linear
 - (xi) SEQUENCE DESCRIPTION: SEQ 10 80:96:
 - Thy Cys Gly Gly Thr Ais Ais Gin Als Als Vai Val Ary Phe Gin Gld 1 10 15

Ale Ale Asn bys Gln tys Gln Slu beu Asp Glu 26 25

- (2) INFORMATION FOR SEC 15 NO. 97;
 - (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 amino soids (B) TYPE: amino acid
 - (C) STRANDBLANESS: Single
 - (D) TOPOLOGY: Linear
 - (NL) SECTENCE DESCRIPTION: SEC ID NO:97:

Sly Cys Gly Ala Asn Lye Gln Lys Gln Gla Lea Asp Slc Ile Ser Ter 1 5 10

Ash Fie Arg Gin Ale Gly Val Gin Tyr Ser Arg

- (1) IMPORPEATION FOR SEQ 10 NO:98:
 - (i) SEQUENCE CHARACTERISTICS: (A) LENGTR: 28 amino acids (B) TYPE: omino acid

123

| 101 | STRANSHIDNESS: | single |
|-----|----------------|--------|
| | | |

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

Sly Cys Sly lie Arg Sin Ale Siy Val Sin Tyr Ser Arg Ale Asp Sin

Glo Gin Gin Sin Als Leu Ser Ser Gin Met Gly Phe 20 25

(2) INFORMATION FOR SEQ 10 86:99:

- (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 507 base pairs
 - (B) TYPE: nucleic acid (C) STRANDEDNESS: pingle
 - (5) TOPOLOGY: linear

(%1) SEQUENCE DESCRIPTION: SEQ IS NO: 99:

| ntgaagatgg | TGAAATOGAT | CCCCGCACGT | CYCLCCGCCG | CSCCFGCAAT | CERCOCOCOCA | 6 |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| Resessors | TGACTFCGAT | CATOGCTOGC | seccesares | TATACCAGAT | GCAGCOGGTC | 3.28 |
| POTTOGRACE | CCCCACTGCC | GTTGGACCCO | GCATCCGCCC | CTSACGTCCC | GACOSCOSÇO | 188 |
| CASTTGACCA | GOCTIGOTUAA | CARCCTOSIC | GATCCCAAOG | TETEOTETEC | GAACAAGGGC | 24 |
| MITCTORTOR | AGGCCGGCCAT | ORGGGGGGACC | GAGGGGGGGGA | TOGOGGACCA | CAAGCTGRAG | 308 |
| AGRECERCOS | ACCADOGGGA | TCTGCCGCTG | restrances | TOACCIANCAT | COACCCCCCC | 368 |
| CCCCCCGGT T | CSOCCACCGC | COACCITICC | GYCYCGGGYYC | CURACCTUTO | GTCGCCGGTC: | 420 |
| NOGCAGAROG | TCACGTTCOT | GAATCAAGGC | GOCTGGATUC | TOTCACGCGC | ATCOCCGATG | 480 |
| SAGITUCTEC | WGGCOGCAGG | CAACTOA | | | | 507 |
| | | | | | | |

- (2) INFORMATION FOR SEQ IS NO. 100:
 - (1) SEQUENCE CRARACTERISTICS:
 - (A) LENGTH: 168 amine acids (B) TYPE: amine acid

 - (C) STRAMDEDNESS: simple (D) TOPOLOGY: linear
 - (wi) SEQUENCE DESCRIPTION: SEO IS NO: 100:

124

| Met. | Lys | Met | Val | liys 5 | Ser | lie | Ala | Alw | Giy 10 | i.au | Thr | AJs | A) a | Ala 15 | Ala |
|------------|------------|-----------|-------------|------------|------------|-------------|------------|------------|-----------|------------|------------|-------------|------------|-----------|------------|
| ile. | Gly | Ala | A.i.s 20 | Ala | Ala | 617 | Val | 7hx 25 | Ser | 114 | Mas | Aia | Gly 30 | GLy | Pres |
| Va1 | Val | Tyr 35 | Gin | Met | Sin | Pro | Val 40 | Val | Phe | 019 | Als | \$2.0 45 | Letu | Fina | Lass |
| Asp | Pro 50 | Ala | Sex | Ala | Pio | 85 85 | Vol | Pro | The | Als | Ala | Gin | 1,00 | The | Ser |
| Les 65 | bea | Ass | Ser | \$mo | Ala 70 | Asp | Pro | Aso | Val | Ser 75 | Pho | Ala | Acri | Lys | 915 80 |
| 342 | læu | Val | Glu | 61 y 85 | Gly | Tle | Gly | Gly | 7hx 90 | G1 a | Ale | Arg | 234 | Ale 95 | Азр |
| 818 | Lys | Leu | Lys 105 | Lys | Ala | Ala | Glu | Dis 195 | Sly | Asp | Lau | Pro | leu 116 | Ser | Phe |
| 508 | Val | The | Asc | 11# | sin. | 220 | Ala 120 | ALs | A18 | Gly | Ser | Ala 125 | The | Ala | quá |
| Val | Ser 130 | Val. | Ser | Gly | Pro | 1.ye 135 | Lou | Ser | Ser | Pro | Val 140 | Thr | Glo | Asn | Val |
| Thr 145 | Phe | Val | Asn | Gin | Gly 150 | Gly | Teys | Mot. | bou | Ser 155 | hrg | Ala | Sex | Ala | Met 160 |
| GIU | Lev | Les | Sla | Ala 165 | Ala | Gly | Aso | | | | | | | | |

(i) SEQUENCE CHARACTERISTICS: (C) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID 80:101:

| COTGGCAATG | TOGTTGACCA | TOGGGGGGGGG | COTOGOCTOS | GCAGATODOG | TGOACGCGGT | 4 |
|-------------|-------------|-------------|-------------|------------|-------------|-----|
| CATTAACACC | ACCORCANT | ACCOUNTAGET | AGTAGCTGCG | CTCAACGOGA | COBATOCOGG | 120 |
| CHICTOCOGCA | CAGTTCAACG | CUTCACCOUT | COCCAGTOL | TATTTGCGCA | DESCRIPTION | 386 |
| COCACCECCA | CCTCAGGGGG | CEGCCATOCC. | CGCGCAA#TG | CARGCTOTEC | COGOGGGGGGG | 245 |
| ACAGTACATO | GOOCTTOTOS | ASTOGGTTGC | cescriconso | AACAACTATT | AAGCCCATGC | 300 |
| S0000CEATE | COSCISACION | CONTRATOR | CSSCOTTAGG | CCAGATTGCC | coordonce | 360 |
| ADEGGCCECA | TODOROGACO | COSCATOSTO | GCCGGGGGTA | GOCCAGATTS | CCCCCCCCCCC | 420 |
| CHACGOGCCG | CATCTCGTGC | CUAATTCCTG | CAGCCCGGGG | GATGUACTAG | TTCTAGAGOS | 430 |

125

| GCCC | CCAC | ce a | GGTG | CACC | 1 | | | | | | | | | | | | 500 |
|------|------------|-----------|--------------------|------------------------|-------------------------------------|--------------------|--------------------|------|-----------|-----------|------|-----------|-----------|-----------|-----------|-----------|-----|
| (2) | INFO | RMNT | 308 | POR | GEQ | ID B | onic | 2: | | | | | | | | | |
| | (\$ } | 130 | I TY | note Pe: Rans | ARAC : 96 amic ECNE GY: | ami o ac 35: | no e id sing | olds | | | | | | | | | |
| | (xi) | | | | | | | | | | | | | | | | |
| | Val. | Bla | Met. | Ber | 5 | The | Va; | 638 | Ala | Giy 10 | Val. | Ala | Ser | Ala | Asp 13 | Pro | |
| | Val | Asp | Als | Val. 20 | 11.0 | Asn | The | The | Cys 25 | Asur | Tyr | gly | Gin | val 30 | Val | Ala | |
| | Ala | Leu | Ann 38 | Ala | The | Aug | Pro | G19 | Ria | 818 | Ali | Gin | 766 45 | Asn | Ala | Ser. | |
| | 220 | Val 50 | Ala | Gla | Ser | Tyr | Leu 55 | Arg | asn | Pho | Les | Ala 60 | Ala | Per | Pro | Pro | |
| | 61 n 65 | Asg | Ala | Ala | Met | #3a 76 | Ala | Gla | i.eu | Gla | A15 | Val | rnd | Gly | Als | Ala 80 | |
| | Gln | Tyr | lle | Gly | 100 85 | Val | Glu | Sex | Val | Ala 90 | ora | Sex | Cys | Ass | 580 95 | Tyr | |
| (2) | TNEC |) ROSA : | ION | FOR | 380 | 10 1 | 00:1 | 332 | | | | | | | | | |
| | (1) | (B) | 1.83 177 373 | IGTH PE: 0 KANDA | ARACT 154 Sonle SONES | i bas | e pa cid | RIXS | | | | | | | | | |

(xf) SEQUENCE DESCRIPTION: 880 10 NO:103:

| ATGACAGAGC | agcagtogaa | PTTCCCCGGGT | ATOMAGGGGG | CEXECUAAGESE | AATCCAGGGA | 60 |
|-------------------|------------|-------------|------------|--------------|------------|-----|
| AATSTCACST | CCATTGATTC | COTCOTTOAC | GAGGGGAAGC | AGTOCOTOAC | CAAGCTCCCA | 120 |
| occoccioss | OCSSTAGGS6 | TTOGGAAGCG | TACC | | | 154 |

(2) INFORMATION FOR SEQ 10 MO:106:

| 2. | MEGUE | ENCE CHARACTERISTICS | 2 |
|----|--------|----------------------|--------|
| | (A) | LEMSTR: 51 amino ac | 3 (38) |
| | 1.19 (| TYPE: smino acid | |
| | (C) | STRANDEDNESS: singl | 8 |
| | | | |

(D) TOPOLOGY: linear

126

| | (x1) | SEQ | USINC | e de | CRI. | rtio | N: S | EQ 1 | 0 89 | : 2.04 | ÷ | | | | | | |
|------|--------|-----------|-----------|-----------------------|---------------------|-------------------------|----------------------|-----------|-----------|--------|-------|---------|-----------|-----------|-----------|-----|------|
| | Mot | The | Glu | Gin | Gin 5 | Trp | Asn | Phe | Als | Gly | 116 | Glu | Ala | Ala | Rla 15 | Ser | |
| | Ala | He | Gin | Gly 20 | Asn | Val | Thr | ser | 11e 25 | 818 | Ser | Seu | Leu | Asp 30 | Gly | Gly | |
| | liye | Uln | Ser 35 | Leu | The | Lys | Lep | Ala 40 | Ala | Als | Trp | Sly | GLy 45 | Ser | GLY | ser | |
| | Glas | Ala S0 | Tyr | | | | | | | | | | | | | | |
| (2) | INSCR | ROSS T | LON | ros s | OBC : | ED NO | 3:10 | 51 | | | | | | | | | |
| | 121 | 18 | TY | KANDE | 28: TOTAL | eic a eic a es: i | se p scid ring | eizs | | | | | | | | | |
| | (81) | | | | | | | | | | | | | | | | |
| | CGCGG | | | | | | | | | | | | | | | | 60 |
| | (0000E | | | | | | | | | | | | | | | | 1.26 |
| | CAAAC | | | | | | | | | | | | | | | | 189 |
| SACA | ATTCE | NA CK | MAGICK | CGCC | TAC | ANX | MAG | TOGG | Fact | WA 7 | VITCE | SINC GE | 19 T3 | rere | GTCS | : | 240 |
| ACCT | Grere | 96 70 | TGNX | RECON | CAK | SIAAC | KKKG | Tact | YXXX | G7 (| 78 | | | | | | 28.2 |
| (2) | INFOR | MAT | ON 3 | OR 9 | SQ 7 | D NO. | :10 | 6-: | | | | | | | | | |
| | 11) | (N) | TXI | M: 4 M: 4 MANUE | 305 unle 0300 | | se ; cid | mārs | i. | | | | | | | | |
| | (ixi | SEQU | ENCE | 983 | CRIP | 1106 | H 88 | 9Q 2D | NO: | 106: | | | | | | | |
| GATO | GTACC | s gi | GOGA | GT00 | 1900 | agez | 617 | TONS | MATG | can o | FECT | cors | * C | T900 | MOAT | | 30 |

GODATACCUA GAGATGTTGG UGGUGGUGGU TUACACCUTG CAGAGKAYCU GTUCIMUUAC

TOTGGCTAGO AATGCCGCTG CGGCGGCCCC GACGACTGGG GTGGTGCCCC CCGCTGCCGA

320

180

127

TERRETISTOR GOSCITERUTS OFFICECACTY CONCECNOR GOOGGETST AFCRETONY 205 GAGUGUTCEG GCTECTECGA TECATGACCA GTTCGTEGCC ACCCTTGCCA GCASCGCCAC 300 CTYSTATURG GOVACTGARG TOGGCRATGE SUCBONSSCC AGCTRAGGER GUARDAGTES 360 SCACGAGAAA CCACGAGAAA TAGGGACACG TAATSCTGGA TTTCCGGGGCG TTACCACCGG 420 AGATCAACTO CUCGAGGATO PACGOOGGOO COGUTTUGGO CTOGOTGOTG GCOGOGGCTO 480 ASATSTSOCA CASCSTSGCS ACTUACCTGT TFTCGGCCGC STUGUESTTT CASTCGGTGG 540 TETEGGGTCT GACGGGGG PEGTGGATAG GTTCGTCGGC GGGTCTGATG GTGGCGCCGG 600 CUTCUCCUTA TETEGORISE ATGROCOTCA CUCCOGGGGGA CUCCGAGGTS ACCGCCGCCC 666 AGETCOGGET TECTOCOGCO GETTACGAGA OSGCGTATGG GCTGAGGETG CCCCCGCCGG TGATCGCCGA GAACCGTGCT GAACTGATGA TTCTGATAGC GACCAACCTC TTGGGGGCAAA 780 ACACCCCGGC GATCGCGGTC AACGAGGCCG AATACGGGG GATGTGGGCC CAACACCCCG 840 COSOSATOTT TOGOTACCCO GAGGGGACGU CGAGGGGGAC GAGGALUETU CTECCOTTAT 900 AGGAGGGGGC GGAGATGAGC AGCGCGGGTG GGCTCCTOGA GCAGGCCGGC GCGGTCGAGG 950 AGGCCTCCGA CACCGCCGCC GCGARCCAST TGATGARCAA TGTGCCCCAG GCGCTCCAAC 3020 AGETYGGOCCA GCCCACGCAG GGCACCACGC CTTCTTCCAA GCTEGGTGGC CTCTGGAAGA 1080 COSTCTOGGC GCATCGGTGG CCGATCRUCK ACATGGTGTC GATGGCCRAC AACCACATGT 1340 CCATGACCAA CYCGGGTGTG TCCATGACCA ACECCTTGAG CFCGATGTTS AAGGGCTTTO 1200 CTYXXXXXXXXX DAXCEGCCTAG PROGTGCAAA CXXXXXXXX AAACGGGGTC CXXXXXXXXXX 1260 SCIUSCISSE CAGOTODOIS CETTOTICES GIVITGOGUGE ISSUETEGO: COCAACTICE 1320 GTCGGGCGGC CTCGGTCGGT TCGTTGTCGC TCCCGCAGGC CTGGGCCGCG GCCRACCAGG 3380 CASTCACCCC GGCGCCGCGC CCCCCCCCCC TGACCAGCCT GACCAGCCCC GCGGAAAGAG 2440 GGCCCGGGCA GATGCTGEGC GGCCTGCCGG TGGGGCRGAT GGCCGCCAGG GCCGGTGGTG 1900 OCCUPACION TOTOCOCCI CITCUSCOSC CACCUPATOT CACCOCCAT TOTOCCCOCCA 1.560 COBOCTAGGA GAGGGGGGG AGACTGTCGT TATTTGACCA GTGATCGGCG GTCTCGGTGF 1625 TTCCGCGGCC GGCTAIGACA ACAGYCAATG TGCATGACAA GTTAGAGGTA TTAGGTCCAG 1880 STITCARCARG GAGAGAGUCA ACRIGGOCTO ACSTITITATO ACSCRIBOCOC ACSCRIBOCO OGACATOCCC GCCCGTTTYS AGGTCCACGC CCAGACGGTO GAGGACGAGG CTCCCCCGGAT 1890 STOSSESTED COCCAMANACA TITOSSESSES SUBCTROMOT GOVATIGACE AGGOGRAPACE 1860 SCIAGACACC ATGGCCCAGA IGAATCAGGC GITTEGCAAC ATGGIGAACA INCIRCLAGGC 1920 SUIGOGICAC GGGCIGGTTO GCGACGCCAA CARCIACHAG CAGCAACAGO AGECTITOCA 1980 SCAGATUCTO ASCAGOTARO GYCAGOTGOT GCAGURCRAT ACTITIACAR GCGRAGGAGA 2040

| ACAGGTTCGA | TGACCATCAA | CTATCAATTC | GOGGATGTCG | ACCCTCACGG | CGCCATGATC | 23.00 |
|------------|-------------|-------------|--------------|-------------|-------------|-------|
| CGCGCTCAGC | CCGCGTTCCT | GGAGGCCGAG | CATCAGGCCA | TCATTCSTGA | PSTSTTGACC | 2160 |
| GCGAGTGACT | TTTOGGGCGG | COCCOGTTON | GCGGCCTYGCE | ASSISTICAT | TACYCACYTG | 5520 |
| GGCCGTAACT | TOURGETGAT | CYACGAGCAG | GOCANDISCOS | ACGGGCAGAA | COTCOAGGCT | 2280 |
| GCCGGCAACA | ACATOSCICA | AACCGACAGC | GCCGTCGCCT | CCASCTGGGC | CTORCACCAG | 2340 |
| CCCAAGGCCA | GGGACGYGGT | GTACGAGTGA | AGTICCIONO | GTGATCCTYC | NOSTGGCAGT | 2400 |
| CTAAGTGGTC | AGTGCTGGGG | ROTTGGTGGT | PISCIECTES | COGGGTTCTT | connector | 2460 |
| CAGTGCTGCT | COGGCT/XGGS | TGAGGACCTC | GAGGCCCXGG | TAGOGGGGGTC | CTTCGATCCA | 2520 |
| TTOSTOSTST | TOTTCGGOGA | 00MOGGETEU | GACGAGGCGG | ATCATCGAGG | CUCOUTCOCC | 2580 |
| GAAGATGCCC | AUGACGTOOG | Tressective | TACCTCTCGG | TTOAGGCXTFT | CCYCCGGGTT | 2840 |
| GTTGGACCAG | ATTEMPORECE | AGATOTSOTT | OCCCAMOSOC | GTGAACGOCA | CCAGGTOGGT | 2700 |
| ocsaveasno | TOGAGGTGCT | COCCACCCC | GGGGAGTTYG | TCGGTCAGAG | CGTCGAGTAC | 2750 |
| CCGATCATAT | TGGGCAACAA | CTGATTCGGC | 9TCGGGCTG4 | TOSTBOATOR | ACTOCÁGCAS | 2820 |
| GGTGCGCACC | CAOGGCCAGG | AGGGCTTCCG | esteechacc | ATCAGATTGG | CTGCGTAGTY: | 2884 |
| GGTTCTGCAG | COCTOCCAGG | COGCTGCGGG | CASSISTESCE | COGATOGCOS | COACCAGGCC | 2940 |
| áscurggges | TOUCTOUTGA | CCAGCGCGAC | COCGGACAGG | ODSCOOGGGA | CONGGTCGCO | 3000 |
| GAAGAACGOC | RGCCAGCCGG | coccarcore | GCCGGARGOTTC | ACCTOGATOC | CCAGGATC | 3058 |

(2) INFURMATION FOR SEQ 10 NO:187:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 191 amino acids (B) TYPE: amino acid

 - (C) STRANOSIMESS: single
 - 100 TOPOLOGY: linear
- (ai) SEQUENCE DESCRIPTION: SEQ ID NO:101:
- Met Val Asp Phe Gly Ala Leo Pro Pro Glo Ile Asn Ser Ala Ary Net
- Tyr Ala Gly Pro Gly Ser Ala Ser Leu Val Ala Ala Ala Gla Mat Trp 25 25 70
- Asp Ser Val Ale Ser Asp Leo Phe Ser Ala Ria Ser Ala Phe Gla Ser
- Val Val Trp Gly Leu Thr Val Gly Ser Trp Ile Gly Ser Ser Ala Gly
- Let Mot Val Ala Ala Ala Ser Pro Tyr Val Ala Tro Met Ser Val The

26 75 Als Gly Gly Ais Glo Lou Thr Ais Als Gla Val Ary Vel Als Als als Ala Tyr Glo Thr Ala Tyr Gly Leo Thr Vai Pro Pro Pro Val Lie Ala Glu Asn Arg Ala Clu Leu Met Ile Leu Ile Ala Thr Asn Len Leu Gly Gin Asn The Fre Ala Ile Ala Vai Asn Glo Ala Clo Tyr Gly Glo Met 136 145 Try Ale Gin Amp Ale Ale Ale Set the Gly Tyr Ale Ale ale Thr Ale 145 150 160 The Ala The Alm The Lew Lew Pro Phe Gin Glo Ala Pro Glo Met The Ser Ala Ciy Gly Leo Leo Clo Gin Ala Ala Ala Val Glo Glo Ala Ser Asp Thr Ale Ale Ale Aso Gin Leu Met Aso Ale Val Fro Glo Ale Leo 195 205 Gin Gin Leu Ala Gin Pro Thr Gin Giy Thr Thr Pro Ser Sec Lys Leo 210 215 226 Shy Gly Leo Tip Lyo The Vel Ser Peo Ris Ang Ser Pro lie Ser Ann Not that Ser Met Ale Axe Am His Met See Wet The Ace Ser Giy Val 245 250 Ser Met Thr Asn Thr Leu Ser Ser Met Lou Lys Gly Phe Als Pro Als Ala Ala Ala Gla Ala Val Gla Thr Ria Ala Gla Asa Gly Val Arg Ala Met Ser ber ber Gly Ser Ser Lea Gly Ser Ser Gly Lea Gly Gly Gly 298 298 300 Vai Ala Ala Aso Leo Sly Arg Ala Ala Ser Val Gly Ser Lon Ser Val 305 310 315 320 Fro Gln Afa Trp Ala Ala Ala Ass Gln Ala Vel The Fro Ala Aia Ary 325Ala Leo Pro Leo Thr Ser Leo Thr Ser Ala Ala Glo Arq Gly Pro Gly 340 545 Cln Met Law Gly Gly Lew Pro Val Gly Sin Met Gly Ala Arg Ala Gly YAS Gly Gly Leu Ser Gly Val Lau Arg Val Pro Pro Arg Pra Tyr Val Met 380 Pro His Ser Pro Ale Ale Sly

(2) INFORMATION FOR BEQ ID NO: 108:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTE: 1725 base pairs

(B) TYPE: oucleic acto (C) STRANDEDNESS: single

(0) TOPOLOGY: linear

(xi) SHOURMEN DESCRIPTION: SEQ ID NO: 108:

| GACGECAGCA | coopposited | MGGG/TYGGA/ | CORGGROOM | The Carrier | GGTCAAGGTG | 8.6 |
|-------------|-------------|-------------|-------------|-------------|---------------|------|
| | | | | | AGTGCAACTA | |
| | | | | | | 120 |
| ATTTCGTTGA | ACTSCCTRCS | AGGTATAGGA | CTTCACGATT | GGTTAATGTA | GCGTTCACCC | 380 |
| COTOTTORSO | PCGATTTGGC | COSACCAGYC | GTCACCAACG | crrecered | SCSCCAGGOS | 240 |
| oscuatorsa | TOSCTTGACT | ACCARTCART | CTTGAGCTCC | CONGCCGATG | CTOGGGGTAA | 300 |
| ATGAGGAGGA | GCACGCGTGT | CTTTCACTGC | GCMACCGGAG | ATMITTGGGG | CCGCGGCTGG | 360 |
| CGANCTTOGT | PCCCTGGGGG | CRACGCTGRA | GECTAGCAAT | GCCGCCGCAG | COSTGCCGAC | 420 |
| GACIGGOGIG | GFGCCCCCGG | CYGCUGACGA | GGTGTCSCTG | CTGCTTGCCA | CACAATTUCG | 480 |
| TACGCATGCG | OCCINCOTATO | AGACGGCCAG | COCCAAGGCC | 9000TGATCC | AYGAGCAGTT | 340 |
| TGTGACCACG | CYGGCCACCA | GOGUTAGTTC | ATATGCGGAC | ACCGAGGCCG | CCAACGCTST | 690 |
| GGTCACCGGC | TARCTGACCT | GACGGTATTC | GROCCGARGO | ATTATEGAAS | TOUTGGATTY | 960 |
| CGGGGGGTTA | CCACCSGAGA | TOAACTCOSC | GAGGATGTAC | GCCGGCCCGG | GTTCGGCCTC | 720 |
| GCTGGTGGCC | GCCGCGAAGA | TOTGOGACAG | CGTGGCGAGT | GACCTGTTTT | COCCEOUTC | 786 |
| GGCCTTTCAG | TOBBIRGICE | 0000TCTGAC | GGTGGGGTCG | TOGATACKIT | OGTOGGGGGG | 840 |
| TOTGATGGCG | GCGGCGGCCT | CCCCCTATCT | CGCGTGGATG | AGCG/PCACCG | CUGGIGGALICIC | 960 |
| CCAGCTGACC | SCOSCOCAGE | recognities | TGCGGGCGGCC | TACGAGACAG | COTATAGGCT | 960 |
| GACCESTCCCC | CUGCUGGTGA | TOGOCGAGAX | COSTACCONA | CHEATGACUC | TGACCGCGAC | 1020 |
| CAACCTCTTG | GEOCÁAAACA | COCCOCCAT | CONSOCCANY | CAGGOOGGAY | MONGCOAGAY | 1980 |
| GTGGGGGGAA | GACGCXGAGG | OGATGTATGG | CTACGCCGCC | ACGGCGGGGA | COGCUAÇOGA | 1140 |
| COCSTTGCTG | CCGTTCGAGG | ACCCCCCACY | GATCACCAAC | CGGGGGGGGGC | TOOTTGABOA | 1200 |
| 3600646609 | OTODAGGAGG | CCATCGACAC | cocooosscs | ARCCARTTGA | TGAACAATUT | 1260 |
| SCCCCAASCS | CTGCAACAGC | TECCCUACCE | MOCSCAGOGO | GTUGTACCTT | CTTCCAACCT | 1320 |
| 90GTGGGC75 | TGGACGGCGG | TCFCGCCGCA | TOTATORCO | CTCAGCAACG | TCAGTTCGAT | 1380 |
| AGCCAACAAC | CACATGTCGA | TRATEGGCAC | gggrgrgrgg | ATGAUCAACA | COTTOCACTO | 1440 |

| | | | 1.51 | | | |
|--------------|-------------|--------------|-------------|-------------|------------|------|
| GATGTTGANG (| GCTEAGCTC | COGCOGCOGC | TENGGOOGTG | GAAACCOCKI | CUGAAAACGG | 1500 |
| GGTCTGGGCG 3 | ASSAUCTOOC | TGGGCAGCCA | CCTGCGTTCG | resersecty | CTTCGGGTCT | 1568 |
| GGGCGCTGGG (| товсскоста | ACTTGGGTGG | GCCGCCCCCC | STORGTTOST | TGTCGGTGCC | 1626 |
| SCCASCATES C | SCOSCOSCOA | ACCAGGGGGT | caddocossos | GCCCGGGCCGC | TOCCOCTUAC | 1680 |
| CAGCCTGACC A | eccecosce | AAACCGCCCC | OGGACACATG | CTGGG | | 1725 |
| (2) IMPORMAT | TION FOR SI | 00 10 NO:10 | 9: | | | |
| \$# \$E |) LENGTH: | | | | | |
| (xi) 886 | vence dedo | CRIPTION: SE | % ID 86:10: | ₹5 | | |

Val Val Asp Phe Gly Ale Leu Pro Pro Gio Ile Ass Ser Ale Arg Met i Tyr Ale Sly Pro Sly Ser Ale Ser Let Vol Ale Ale Ale Lys Met Trp Amp Ser Val Ala Ser Amp Leu Phe Ser Ala Ala Ser Ala Phe Gin Ser 35 46 45Val Val Trp Gly Len Thr Val Gly Ser Trp fie Gly Ser Ser Ala Gly 50 60 Low Met Ala Ale Ale Ale Ser Pro Tyr Val Ale Trp Het Ser Val Thr 65 70 75 86 Ala Gly Gln Ala Gln Leo Thr Ala Ala Gln Val Arg Val Ala Ala Ala Ale Tyr Giu The Ale Tyr Arc Leu Thr Val Fro Fro Pro Val Ile Ale Gib Asn Ary Thr Gis Lew Met Thr Lew Thr Ale The Asn Lew Lew Gly Gin Ann the Pro Alé 116 Sio Ale Ann Sin Ale Ain Tyr Sex Gin Met 130 $\,$ 140 Trp Gly Gld Asp Ais Glu Ala Het Tyr Gly Tyr Als Als Thr 81% Ais 145 The Als The Glu Ala Leu Leu Pro The Glu Asp Ala Pro Leu Fle The Ann Pro Gly Gly imp Leu Glu Glo Ale Vel Ale Vel Glo Ale Ile 180 | 180 | Asp The Ala Ale Ale Asn Gin Led Met Asn Asn Val Pro Gin Ale Lou 200

| Oin | 91n 210 | 1/913 | Als | GLa | Pro | Ala 215 | | (C) y | Val | Val. | 220 220 | | 545 | loys | Letin |
|------------|------------|------------|------------|------------|------------|---------------------------------------|------------|-------------|------------|--------------|------------|-------------|------------|------------|------------|
| 225 225 | GLy | Lou | Trp | Tita | A1a 230 | Val | Ser | 800 | ная | Leto 2.85 | Sar | Px'o | Les | Sec | Asn 240 |
| Yal | Ser | Ser | 110 | A1s 245 | Asn | Asti | His | Merit | Ser 250 | Mes | Most | Gly | The | Gly 255 | Va) |
| Ser | Mer | The | Asn 266 | Thr | Leu | Ris | Ser | Met. 265 | ken | Lys | Gly | 200 | Ala 270 | fro | Ala |
| Ala | Ala | 610 275 | Alk | Val | Gla | The | A18 289 | Als | Glu | Asn | Sly | Val. 285 | Trp | Ala | Mes |
| Ser | Ser 290 | Leu | Gly | Set | Gln | 1.95 295 | Gly | Ser | Ser | Leu | Gly 300 | Ser | Sec | Gly | Len |
| GLy 305 | Ala | Gly | Val | hia | Ala 310 | Agn | Leu | Gly | Arg | A18 | Alu | Ser | Va) | Gly | Ser 320 |
| Leu | Ser | Val | 220 | 2ro 325 | Ala | Tro | Ala | Alla | Ala 330 | Ass | Sin | Ala | Val | Thr 335 | Pro |
| Ala | Ala | Arg | Ala 340 | Legis | Pro | Leu | The | Ser 345 | Leys | Thr | 362 | Ala | Ala 350 | Sin | Thr |
| Ala | Ero | dly 355 | His | Met. | Lena | Gly | | | | | | | | | |
| THEFT | 7N8.5 17 | rent r | 330 3 | 2002 | COL AND | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 | 57 | | | | | | | | |

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3027 bess pairs
(B) TYPE: nucleic scid
(C) STRANDERMESS: single
(D) TOPOLOGY: linear

(MI) SEQUENCE DESCRIPTION: SEQ IS NO:116:

| AGTTCAGTCG | acantertac | YGACGGRICITE | TATCCACGAT | GGCTGAGACA | ACCHARCCAC | 6 |
|-------------|-------------|--------------|------------|------------|-------------|-----|
| CCTCGGACGC | 9996ACA900 | CANGCOGACG | CGA1GGCG7T | 9900000GAA | GCCGAAGCCS | 1.2 |
| OOGAAGOOGA | AGCGCTGGGC | 0000000000 | GGGCCCGTGC | COGTGOOGOG | CUSTIGAAGE | 18: |
| GEGAGGCGCT | SGCSATGGC | CCAGGGGGAGG | ACGÁGAACCT | CCCCGAGGAT | ATGCAGACTG | 24 |
| GGAAGAGSCC | GARGACTATO | achactatga | CGACTATGAG | GCCGCAGACC | AGGAGGCCGC | 304 |
| ACGGTCGGCA | TCCTGGCGAC | SECRETTECO | OCTG099TTA | COAAGACTOT | CCACGATTGC | 364 |
| CATRICCOGCO | GCAGTCGTCX | TOATOTRORG | CTTCACCGGG | CTCAGOGGAT | ACATTOTOTO | 42 |
| GCAACAOCAT | GREGOCIACOS | AACGCCAGCA | 9000600000 | GUSTTURCOG | COGGRAGOCAA | 488 |
| CCAAGGTGTC | ATCAACATGA | CCTCGCTSGA | CTTCAACAAG | GCCAAARAAG | ACGTCGCGCG | 545 |

PCT/US97/18293 -

| TOTGATOGAC | AGUTOCACCO | GOGAATTCAG | GGATGACTYC | CAGCAGGGG | CASCOGATTY | 600 |
|-------------|------------|-------------|--------------|-------------|-------------|-------|
| CACCAAGGTT | GTCQAACAGY | CCMAAGFEGT | CACXXXXXGGC | ACSGEVAACE | CGACACCUST | 660 |
| CGAATOCATO | AACGAGCATT | cescenter | GCTCGTCGCG | 6CGACTICAC | GGGTCACCAA | 720 |
| TTCCGCTGGG | GCGAAAGACG | AACCACGTGC | GYGGCGGGTC | AAAGTGACCG | TGACCGAAGA | 280 |
| SUGGGGACAS | TACAMMATET | CGAAAGTTOA | STICGIACCO | TGACCGATGA | COTACGCOAC | 840 |
| GTOAACADÇG | ASSCRETGA | COCCACOGAA | GTCGCTGAGA | TEGACTEAGE | CGCAGGCSAA | 900 |
| GCCGCTGATT | OGGOGACOGA | GCCATTTGAC | ACCGACTOTO | CARCGGRATC | TACCUCGCAG | 960 |
| AAGGGTCAGC | GGCACCGTGA | OCTGTGGCSA | ATGCAGGTEA | COTTONARCO | CSTTCCGGTS | 1020 |
| ATTOYCATOO | TGCTCATGTT | GATCTCTGGG | GGCGCGACGG | GATGGCTATA | CCTTGAGCAA | 1.080 |
| TACSACCESA | TCAGCAGAGE | GACTOCGGCG | ocsecceroc | TGCCGTCGCC | GCGGCGTCTG | 1140 |
| ACCGGGAGAAT | CHCGCTGTTG | TGTATTCACC | CGACACGTOG | ACCAMGACTT | CCCTACCOCC | 1200 |
| AGGTCGCACC | TOGOCOGOGA | TETROCTOTO | TATACCACCA | GTTCACGCAG | CAGATOSTGG | 1260 |
| CTCCGGCGGC | CAAACAGAAG | TCACTGAAAA | CCACCCCAA | GGTGGTGCAC | GCGGCCGTGT | 1390 |
| CGGAGCTACA | TOOGGATTOG | SCCSTCSTTC | TGGTYTTGT | CGACCAGAGC | ACTACCAGEA | 1380 |
| ASSACASOCO | CAATCCGFCG | ATGOCGCCA | GCAGCGTGAT | GGTGACCCTA | POTEBBAADOR | 1440 |
| ACCCCAATTG | SCTSATCACC | AAGETCACCC | CGGTTTAGGT | TECCETAGGE | GG7CGCCANG | 1500 |
| TCTGACGGGG | SCSCCGGTGG | crecreered | GAGATACTEG | CCOTTCTCCG | GACAATCACG | 1960 |
| GOOGRACUTC | AAACAGATOT | COCCCCCTGT | CTAATOGGCC | GOGTTATTTA | AGATTAGTTO | 1620 |
| CCACTGTATT | TACCYGATGT | TOMMETOIT | CAGCTGGATT | TAGCTTORGG | GCAGGGGGGC | 1.680 |
| TOGTGCACTT | TOCATCIGGG | GTTGTGACTA | CTTGAGAGAA | TTTGACOTGT | TGCCGACGTT | 1740 |
| GTTTGCTGTC | CATCATTGGI | SCIASTIATO | GCCGAGCGGA | AGGATTATCO | AAGTGGTGGA | 1800 |
| CTTCGOGGGG | PTACCACCGG | AGATCAACIC | CGCGAWATG | TACGCCGGCC | COUNTRESSE | 1866 |
| creacraggs | GCCGCCGCGA | AGATGTGGGA | CAGCGTOXXXX | NOTGACCTOT | TTTGGGCCGC | 1920 |
| GTCGGCGTTT | CAGTCGGTNG | TOTGGGGTCT | GACGACGGGA | TOSTGGATAG | STTCGTC890 | 1980 |
| GGGICTGATO | GT00CGGCGG | CCTOUCOGTA | POTOGOGYSK | ATGREGATCA | CCGCGGGGGCA | 2040 |
| ggoogagong | ACCGCCGCCC | ACCITICOGGT | POLITICATION | CONTACGAÇA | SOTATICKES | 2100 |
| GCTGACGGTG | 0000000000 | TGATCGCCGA | GAACCGTGCT | gabotgatga | TTCTGATAGC | 2160 |
| GACCAACCTC | TTORGGCAAA | ACACCCCOSC | CATOGOGGEO | AACOAGGCCGS | AATACDGGGA | 3226 |
| Satgregece | CAAGACGCCG | COGOGATETY | TRECTACCCC | OCCAÇOGOOS | CGACGCCGAC | 2280 |
| Coassostes | CTGCCGTTCS | AGGACGOCCC | ACTGATCAEC | AACCCC99CG | GGCTCOTTGA | 2349 |
| GCAGGCCGTE | GCGGTCGAGG | AGGCCATCGA | CACCBCCGCG | OCGABCCAOT | TOATGAAGAA | 2450 |
| | | | | | | |

WO 98/16646 PCT/US97/18293

134

| TOTGCCCCAA | GCGCTGCAAC | AACTONUCCA | OCCCACGAAA | AGCATCTGGC | COTTOGACCA | 246 |
|------------|------------|------------|-------------|------------|--------------|------|
| actgagtgaa | CTCTSGAAAG | CONTCTCSCC | GCAFCTGTCG | CCCCTCAGCA | ACATOSTISTC. | 252 |
| GATGCTCAAC | AMECACGEGI | CGATGACCAA | CTCSGGTGTG | TOGATGGCCA | GCACCTTGCR | 358 |
| CTCAATGTTG | AAGGGCTTTG | CTCCGCCCCC | GOCTCAGGCC | GTGGAAACCG | CGGCGCAAAA | 264 |
| CGGGGTCCAG | GCCATGARCT | CONTEGEN | CCAGCTGGGT | TOSTOSCIQU | GTTCTTCGGG | 2706 |
| rerecessor | OCCUTOCCC | CCAACTIGGG | Tessogogoec | POSGTUGGET | COTTCTCGGT | 276 |
| SCOGCAGGGC | TGGGCCGCGG | CCAACCAGGC | GGTCACCCCG | GOGGGGGGGG | COCTGOCGCT | 290 |
| GACCAGCCTG | ACCASCACCG | CCCAAACCOC | CCCCGGACAC | ATGUTOGGGG | COCTACCGCT | 2896 |
| GGGGCAACTG | ACCARTAGÉS | 4Cascacett | COGDEGGGTT | AGCAATGOGT | TOCOGRAPACE | 2946 |
| gecooggees | TACGTAATGC | cocererect | CONCRECCOR | TAACGOCGAT | COGGAGGGAA | 3000 |
| TOCGGGCCCT | CTWYCOGOGC | ACCGATC | | | | 3627 |
| | | | | | | |

- (2) IMPORMATION FOR SEQ ID NG:111:
 - (1) SHOURNCE CHARACTERISTICS:
 - (A) LENGTH: 398 amino solds
 - (8) TYPE: amino scid
 - (C) STRANDEDWESS:
 - (B) TOPOLOGY: Linear
 - INTO SECUENCE DESCRIPTION: SEC 10 NO.1111:
 - Vol Vol Amp Pom Sly Ala Leu Pro Pro Siu Lie Amn Ser Ala Arg Met
 - Tyr Ais Gly Pro Cly Ser Ala Ser Leu Val Ala Ais Ais Lys Met Try 20 30 30
 - Asp Ser Val Ala Ser Asp Less Phe Ser Ala Ala Ser Ala Phe Cln der 35 49
 - Val Val Trp Gly Les Thr Thr Gly Ser Trp lie Gly Ser Ser Ale Gly 50 60
 - Low Met Val Ala Ala Sex Pro Tyr Val Ala Try Met Ser Val Thr 65 75 80
 - Als Gly Gle Als Glu Leu Thr Als Als Glu Val Arg Vel Als Als Als 90
 - Als Tyr Glu Thr Als Tyr Gly Lou Thr Val Pro Pro Pro Vol Ile Ala 198
 - Giu Ann Arg Ala Gio Leo Mot 134 Lou Fle Ala Thr Ann Lou Leo Gly 125 120
 - Glo Ann The tro Ale tie Ala Val Ash Glu Ala Glu Tvr Gly Glu Mot

135

140 Top Ala Sin Asp Ala Ala Ala Set Phe Gly Tyr Ale Ala Thr Ala Ala 145 150 160 The Ala The Gio Ala Leo Leo Pro Phe Glu Asp Ala Pro Leo Ila The Asn Pro Cly Cly Lou Lou Ciu Glo Ale Val Ale Val Ale Glu Ale 11e Asp Thr Six Aid Alm Asn Gin Leu Met Aso Asn Val Pro Gin Aia Leu Glo Glo Lew Als Glo Pro The Lye Ser He Trp Pro Pae Asp Gin Lew Ser Glo Leo Trp Lys Alis Ilv Ser Pro Wis Leo Ser Pro Leo Ser Aso The Val Ser Not Law Ass Ass Mis Vel Ser Met Thr Adm Ser Gly Val 245 256 Ser Met Ala Ser The Leo Sie Ser Met Leo Lys Gly Phe Ala Pro Ala Als Ala Gin Ais Vel Gio Thr Ala Ala Gin Asn Gly Vel Gin Ais Met Ser Ser Lea Gly Ser Gle Lea Gly Ser Ser Lea Sly Ser Ser Gly Lea Cho 295 300 Gly Ala Gly Vel Ala Ala Ash Leu Gly Arg Ala Ala Sor Val Gly Sor Let Ser Vol Pac Gin Ale Top Ale Ale Ale Asc Gin Ale Val The Pro Ala Ala Arq Ale Lou Pro Leo Thr Ser Leo The Ser Ala Ala Gla The Ala Pro Gly Hix Net Leu Gly Cly Leu Pro Leu Gly Gla Leu Thr Azr 365 360 169 Ser Gly Gly She Gly Sky Val Ser Asu Als Leo Arg Met Pro Pro 376 380 Arg Ala Tyr Val Net Pro Arg Val Pro Ala Ala Gly

(2) INFORMATION FOR SEC 10 NO.112:

- (5) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1616 base pairs
 - (B) TYPE: sucletc acid
 - (C) STRANDEDWESS: single

 - (B) TOPOLOGY: libear

[NI] SEQUENCE DESCRIPTION: SEQ ID NG: 112:

| CATCGGAGGG | AGTGATCACC | ATOCTOTOC | ACCOMMISSION | ACCGCAGTAA | ATACCGCACG | \$6 |
|--------------|-------------|---------------|-------------------|--------------|-------------|-------|
| OCTGATOSCC | GBCGCGGGTC | COGCTCCAAT | GCTTGCGGCG | GCCGCGGGAY | GGCAGACGCT | 120 |
| TTCGGCSGCT | CIGGACGCIC | AGGCCGTCGA | OFFGACUSCO | CUCCTGAACT | CTCTGGGAGA | 180 |
| AGCETGGACT | GGAGGTGGCA | GCGACAAXXC | COTTGCGGCT | GEAACGCCGA | TGGTGGTCTG | 245 |
| GCTACAAACC | GCGTCAACAC | AGENCIAAGAC | CCGTGCGATG | CAGGCGACGG | CGCAAGCCGC | 300 |
| SGCATACACO | CAGGCCATOR | CCACGACGCC | GTCOCTGCCG | GAGATOSCCS | COMMODAGAT | 385 |
| CACCCAGGCC | GECCTEAGGG | CCACCAACTT | CTTCGGTATC | AACACKRTCC | GGATCGCGT? | 420 |
| GACCGAGATO | GATTAPFFCA | TOCGTATGTG | GAACCAGGCA | GCCCTGGCAA | TOGAGGTOTA | 480 |
| CCAGGCCGAG | ACCGCGGTTA | ACACGCTTTT | CGAGAAGCTC | CACCCCATCC | CGTCGATCCT | 540 |
| TGATCCCCGC | GCGRGCEAGA | GCACGACGAA | OCCGATETTC | OGANTOCCCT | étectage Au | 600 |
| CTCAACACCG | GTTGGCCAGT | TOCCOCCOC | GGCTACCCAG | ACCCTCOCKC | AACTSGGTGA | 660 |
| CATGAGCGGC | CCGATGCAGC | AGCTGACCCA | GOCCOTTGCAG | CASSTGACOT | CGTTGTTGAG | 3.20 |
| CCAGGTGGGC | GGCACCGGCG | GCGGCAACCC | AGCCGAGGAG | GAAGUOSCOC | AGATGGGCCT | 780 |
| GCTCHHCACC | AGTOCOCTGT | OGRACIOATCC | SCIOGCIGGS | GGATYCAGGGC | CCASCGCGGG | 840 |
| CRCGGGGGCCTG | CTGCGCGCGG | AGTCGCTACC | TGGCGCAGGT | GGGTCGTTGR | CCCGCACGCC | 900 |
| GCYGATGTCT | CAGCTGATEG | AAAAGCCGGT | TOCCCOCTCS | STEATGOODS | COCCTOCTOC | 968 |
| decarestes | GCGACGGGTS | GOGCCGCTCC | agressfoce | GOAGCGAYGG | SCCAGGGTGC | 1626 |
| GCAATCCGGC | GROTOCACCA | seccesarer | GGTCGCGCCCA | GCACCUCTCC | COCAGGAGGG | 1080 |
| TGAAGAAGAC | CACGAGGACG | NOTIGOGACIGA | AGAGGACGAC | TGGTGAGCTC | COGTAATGAC | 1140 |
| AACAGACTIC | COUGCOCACOC | GOGCCCGGGANAG | ACTYCCCAAC | ATTTTGGGGA | GGAAGGTAAA | 1.200 |
| CAGAGAAAGT | ACTCCAGCAT | GGCAGAGATG | AAGACCGAYG | COGCTACCCT | CUCCUCAGDAD | 1260 |
| CCAGGTAATT | TOGAGCGGAT | CTCCGGCGAC | CTGRAAACCC | AGATCGACCA | GGTGGAGTCG | 1326 |
| ACCCCACCTT | CCTTCCAGGG | CONGTESCES | 99080GGGG | GUACODICCIGO | OCREGOGGO. | 1380 |
| STESTSCECT | POCAAGAAGC | rgccaataag | CAGAAGGAGG | AACTEGADSA | GATCTCGACG | 3440 |
| AATATTOOTT | AGSCCGGCGT | CCAATACTCG | AGGGCCCACS | AGGAGCAGCA | GCAGGCGCTG | 1500 |
| TOCTOGUAAA | TECCTTOTE | ACCEGUTAAT | ACGAAAAGAA | ACGGAGCAAA | AACATGACAG | 1560 |
| ACCACCAGEG | GAATTTCCCC | COTATODAGG | CCCCGGCAAG | CGUAATCCAG | CGAAAT | 1616 |
| | | | | | | |

⁽²⁾ INFOTONATION FOR SEQ 10 NO:113:

⁽i) REQUENCE CHARACTERISTICS:

⁽A) LENGTH: \$32 base pairs (%) TYPE: neeleic seid

⁽C) STRANDEDNESS: single

(D) TOPOLOGY: Linear

| 100 5 3 | SCHOOLS | TEXACOTORYON . | 45.000 | YD | Sec. | 473 | |
|---------|---------|----------------|--------|----|------|-----|--|

| CTAGTGGATG | GGACCATGGC | CATTITCEGC | AGTOTCACTG | CCTTCTCTGT | TGACATTTTG | 40 |
|------------|------------|------------|------------|------------|------------|------|
| GCACGCCGGC | GGAAACGAAG | CACTGGGGTC | GAAGRACOGC | FECGCTSCCA | TATUSTCOSS | 120 |
| AGCTTCCATA | COTTOGTGCG | GOOGGRAGAG | CTTSTCGTAG | TOGGOGGGGA | TGACAACCTC | 1.00 |
| TCAGAGTGCS | CICARROGIA | TAAACACGAG | AAAGGGCGAO | ACCUACCOAA | GGTCGAACTC | 240 |
| GCCCGATCCC | CARLLICKA | ATTETACGCG | AACTOGGOGT | TGCCCTATGC | GARCATCOCA | 300 |
| GTGACGTTGG | CTTOGGTCGA | AGCCATTGCC | PGACCEGCTT | COCTGATCGT | COSCOCEAGG | 360 |
| TTCTGCAGCG | COTTOTTCAG | CTCCGTACCC | GTGGCGTCCC | ATTTTTCCTG | GACACCCTGG | 420 |
| POOTOODAY | AA | | | | | 432 |
| | | | | | | |

(2) INFORMATION FOR SEQ ID NO:114:

- (i) SEQUENCY CHARACTERISTICS: (A) LENGTH: 368 amino soids
 - (B) TYPE: amino acid (C) STEAMDEDNESS: single
 - (D) TOPOLOGY: linear
- (xi) SEQUENCE DESCRIPTION: SEO ID MO:116:
- Met Leu Trp Bis Ala Met Pro Pro Glu Xaa Asn Thr Ale Ary Leu Met λ 5 10 18
- Ala Gly Ala Gly Pro Ala Pro Met Leu Ala Ala Ala Ala Gly Trp Gln 20 25 30
- Thr Leu Ser Ale Ale Lou Asp Ale Glo Ale Val Glo Leu Thr Ale Arg 35 43
- tou Asn Ser Leu Gly Giu Aia Trp Ter Gly Sly Gly Ser Asp bys Ala 50 55
- Leu Ala Ala Ala Thr Fro Met Val Val Trp Leu Gie Thr Ala Ser Thr 65 75 ac
- Gin Als Lys The Arg Ala Met Gin Als The Ala Gin Ala Ala Ala Tyr 85 96 95
- The Glo Ala Met Ala The The Pro See Lee Pro Glu 146 Ala Ala Asa 100 105 116
- His liv Thr Sin Ala Val Leu Thr Ala Thr Asm Phe Mie Gly lle Aem 115 120 120

The lie Pro lie Ale Lou Tor Giu Met Asp Tyr Pin fle Arg Met Trp Aso Gio Als Ais loo Als Met Glo Val Tyr Glo Als Gio Tor Als Val Asn for her the Glo bys Lou Glu Pro Met Ala Ser He Leu Asp Pro Oly Ala Ser Gin Ser Tur The Asa Pro lie The Gly Not Pro Ser Pro Gly Ser Ser The Pro Val Gly Gin Low Pro Pro Ata Ala The Gln The Len Gly Sin Len Sly Glo Met Ser Gly Pro Met Gin Cin Leu The Gin Pro Leu Gin Gin Val Thr Ser Leu Phe Sex Gin Val Gly Cly Thr Gly Gly Gly Asn Pro Ala Asp Glo Glu Ala Ala Gin Mar Gly Leo Leo Gly The Ser Pro Lou Ser Ass Ris Pro Lou Ala Cly Cly Ser Gly Pro Ser Als Gly Ala Ciy Leu Leu Arg Ala Glu Ser Lau Pro Gly Ala Gly Gly 275 280 285 Ser ben The Arg The Pro ben Met Ser Cin ben file Glu Lys Pro Val Als Pro Ser Val Met Pro Als Als Als Als Bly Ser Ser Als Thr Gly Gly Als Als Pro Val Gly Als Cly Als Mer Gly Glo Gly Als Glo Ser Gly Gly Ser Thr Arg Pro Gly Len Val Ala Pro Ala Pro Len Ala Gio Glu Ard Glu Glu Asp Asp Glu Asp Asp Trp Asp Glo Glu Asp Asp Trp

(2) IPFORMATION FOR SEQ 15 NO:115:

- (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 100 amino acida
 - (B) TYPE: amino sold
 - (C) STRANDEDNESS:
 - (D) YOPOLOGY: Linear

(M1) SEQUENCE DESCRIPTION: SEO ID MO1115:

Met Ala Ciu Not Lya Thr Aco Ald Ala Thr Leu Ala Cio Clu Ala Gly I

| Asn | Pho | 121 11 | Arg 20 | 1.1-9 | Sez | Gly | Asp | 1.eu 25 | Lys | The | Sin | Ile | Asp 30 | nln | Val. |
|-----------|-----------|-----------|-----------|-------|-----------|------------|-------------|------------|-----------|-----------|------------|------|-----------|-----|-----------|
| Glu | Ser | 7hr 35 | Ala | Gly | Ser | i.eio | G\$11 40 | GLy | 32n | Tre | Arrg | 01y | Ala | Ala | Gly |
| The | Als 50 | Ala | Gla | Ala | Ala | Val. 55 | Val | Arg | Die | QIn | 61 a 60 | діа | ala | Ass | liys |
| Gin 65 | Lys | Glin | Glu | Lou | Asp 70 | Clo | lle | Ser | The | Aen 75 | lle | Arg | Gin | Ala | Gly 80 |
| Val | Gln | Tyr | Ser | Arg | Als | Asp | era | Glu | 63n 90 | Gin | Gla | Alia | Leia | Ser | Ser |
| Gln | | Gly | | | | | | | | | | | | | |

- (3) INFORMATION FOR SEQ ID NO:116:
 - (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 396 base pairs
 - (b) TYPE; nucleto soid
 - (C) GTRANDADNESS: single (D) TOPOLOGY: linear
 - (gi) SEQUENCE OSSCRIPTION: SEQ TO NOTIFE

| COST CACCORD. | HACKITGARARA | CULTAGATUSA | CCAGGTGGAG | TOSACGGCAG | GTTCGTTGCA | 60 |
|---------------|--------------|-------------|------------|------------|------------|-----|
| oosecygard | CGCCGCGCGC | CECEGACGEC | PSCCCARGCC | cocreates | GCT7CCZAGA | 120 |
| AGCAGCCAAT | AAGCAGARGO | AGGAACTOGA | CGAGATOTOS | ACGARTATTC | GTCAGGCCGG | 180 |
| CGTCCAATAC | TOGREGGECEG | ACGAGGASCA | GCAGCAGGCG | CTGTCCTCGC | AAATGGGCTT | 240 |
| CTGACCOSCT | AATACGAAAA | GAAACGGRGC | AAAAACATGA | CAGAGCAGCA | GPOGRATTIC | 300 |
| OCUGGTATCG | ASSCCOCCEC | AAGCGCAXTC | CACGCAAATG | TCACGTCCAT | TONTROCCTC | 360 |
| CPFGAMGAGG | GGAAGCAGTC | CCTGACCAAG | CTOGCA | | | 398 |
| | | | | | | |

- (2) INFORMATION FOR SEC 10 NO:117:
 - (1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 80 amino acids
 - (D) TYPE: amino acid
 - (C) STRANGEDNESS: single
 - (D) YORGLOGY: Linear
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:
 - The Ser Gly Asp Les Lys Thr Gin lie Asp Gla Wal Glo Ser Thr Ala

| WO 98/16646 | PCT/ES97/1829 |
|-------------|---------------|
| | |

140

| \$ | | | | 5 | | | | | 20 | | | | | 35 | | |
|-----------|-------------------|------------------------------------|-----------------------|----------------------|----------------------|---------------------|------------|-----------|-------|-----------|-----------|-----------|------------|---------|-----------|-----|
| 619 | Sec | Seu | Gin 20 | 9) y | Gla | Trp | Arg | Gly 25 | Ala | Ala | Gly | 7902 | A1.a 30 | Als | Gla | |
| ALs | Ala | Va1 35 | Val. | Arg | Pha | 6) n | 61 u 40 | Ala | Als | Asu | Lys | 61a 45 | Lys | Gin | 010 | |
| Len | Asp 50 | Glu | 110 | Sex | Thr | Asn SS | Ile | Arg | Gla | Ala | Gly 60 | Val | Gln | Tye | Ser | |
| Arg 65 | Ala | Asp | Glu | Glu | 01 n 70 | sla | Gla | Als | lou | Sex 75 | Sex | Sko | Met | Oly | Pho 80 | |
| (2) inpu | reat: | con s | eor s | EQ 3 | LD No | 2111 | 3 : | | | | | | | | | |
| (1) | (A) | JENCT LASE TYL 971 TOL | GTH: H: F RANDS | 38° opis CONEL | bes sic a St s | re pr laid | sirs | | | | | | | | | |
| (%i) | | | | | | | | | | | | | | | | |
| GTGGATCC | G AT | rcovo | rori | TXX | CTAT | TOT | ACGC | COSSC | TC C | xocon | 7600 | C TA | TECK | ARCA | i | 60 |
| TOMAGTGA | C G | rrgcc | TTCS | GTC | SAAG | CCA | TTGC | CTGA | ext 6 | GCTT | CGCT | G AT | CSTC | idaac | | 120 |
| CCACKTTCT | VS CA | GCGC | TTTG | Trip. | ASCY | CGG | TAGE | XX 3X | GC 6 | TXXX | ATTI | 7 70 | e rac | acac | | 180 |
| CUTGGTACO | c cr | CCGA | ACCG | CTA | cege | sec | agge | rger | GC G | AGOI | TGGT | C AC | GGA. | rgei | | 240 |
| TOCCUTCUI | C 88 | WW. | CZGAA | TGA | Ares | iacs | TOAC | ATTT | ec e | †GG# | TTGC | G 03 | TGCV: | 734,750 | | 300 |
| CCFCGATAC | e eg | CGAA | APEC | CM | TGCT | GCT | CTGT | CATG | rr r | TIGO | rcos | 7 77 | crrr | rom | | 360 |
| ATTAGCIGG | T CA | GAAG | OCCA | 777 | (50G) | | | | | | | | | | | 397 |
| (2) INFOR | MATT | 08 8 | OR S | EQ 3 | 0 80 | :119 | 0 | | | | | | | | | |
| (2) | (A) (B) (C) | ENCE LAR TYP STR TOP | GTB: E: B ANDE | 273 acle DMDS | bas ic s S: 8 | e pa old ingl | ira | | | | | | | | | |
| (8.4) | SEÇO | SNCE | DRS | CKIP | TION | : SE | Q 10 | 30: | 119: | | | | | | | |

CARCACRAGE ATCTCRETTE UCCCAACGOD GCTGGCDAGE GCTCCGTTGD GGGGGGAGC 60 TECCCCCCG ATSCTTCCTC TOCCCGCAGC CECCCCTGGN TOGATGGACC ACTTCCTACC 120 TYCOCHACET TYCETYCHET GYCTGYGCA TAGCGGYGAC CCCGGCGCGCC ACCTCCGGAG 190 WO 98/16646 PCT/0897/18293

141

| TGTT | 9999 | ec c | AWGO | pess | T OG | 9790 | TTCS | 60 | 00900 | RACG | CAG | ecog. | 107 | 36.40 | ogna | 06 | 249 |
|-------|--|------------|----------------------------------|----------------------|------------------------|--------------------|------------|------|-------|-----------|-----|-------|-----|-------|-----------|-------|-----|
| 6600 | 0000 | TT C | goog. | 6776 | G 08 | TOTT | TGCC | CA. | | | | | | | | | 272 |
| (2) | INFO | BMAT | TON ! | FICK | 032 | 10 10 | 0:12 | 0): | | | | | | | | | |
| | (4) | (E) | Denci LG TY TY TO | NOTH PELI RAND | c 20 amin Stars | ami c ac 63: | no a le | | * | | | | | | | | |
| | ixí; | 980 | JENCE | s per | SCRI | PTIG | N: S | eo o | Ø W | 120 | | | | | | | |
| | Asy | Pro | Val | Asp | Ala S | VA3 | 716 | Asr | Th: | Tn: | Cys | Ass | Tyi | 615 | 915 | : Vai | |
| | Val | Ala | | Leta 20 | | | | | | | | | | | | | |
| (2) | INEU | RMATI | ON F | 02 : | DBQ. | ID B | 0:12 | 1: | | | | | | | | | |
| | and and | (8) (C) | ENCE LES TYP STP TOP | GTH B: / AND! | 1.5 XMA NO CONE: | amia a ec | no a id | | | | | | | | | | |
| | (×i.) | BEQU | RENCE | ORG | CBJ I | 27102 | F (83 | eo i | D NO | :121 | ¢ | | | | | | |
| | Ala | Val | Gla | Sec | 91y 5 | Mac | Leu | Ala | Leri | 6J y | The | Pro | Ala | Pro | 8er 15 | | |
| (2) 2 | NEO | WAT1 | ON I | OR F | (80) | 0 80 | 122 | 2. | | | | | | | | | |
| | (1) SCOURNCE CHARACTERISTICS: (A) LENGTH 10 mmins solde (P) TYPE: mmins sold (C) STUNNCESNESS: (D) TOPOLOGY: Liture: | | | | | | | | | | | | | | | | |
| (| (x1) SEQUENCE DESCRIPTION: SEQ 10 NO. 120: | | | | | | | | | | | | | | | | |
| | Ala 1 | Ala. | Bat : | Lys | eng S | Arg. | Fire | Gly | Asp | 619 10 | Pro | Less | Glu | Ala | Ala 15 | Lys | |
| | (du | 63 y . | Arg | | | | | | | | | | | | | | |

- (2) ISFORMATION FOR DEC 10 NO-123:
 - (1) SEQUENCE CHARACTERISTICS:
 - (A) LEMSTH: 15 amino acida
 - (C) STRANCEONESS:
 - (9) TOPOLOGY: linear
 - (x4) SEQUENCE DESCRIPTION: SEG 10 NO:123:

Tyr Tyr Trp Cys Pro Gly Gin Pro Phe Asp Pro Ala Trp Gly Pro 1 10 10

- (3) INFORMATION FOR SEQ ID NO: 124:
 - (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 14 amine soids (B) TYPE: amino acid
 - (C) STRANGEDMESS:
 - (D) TOFOLOGY: linear
 - (x3) SECUENCE DESCRIPTION: SEQ 10 NO: 124:

Asp Tle Gly Ser Clu Ser Thr Glu Asp Glo Gln XAs Ala Ya)

- (2) INFORMATION FOR SEQ 10 NO.1251
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 13 amino acids
 - (B) TYPE: amino soid
 - (C) STRAMBEDNESS:
 - (0) TOPOLOGY: Linear
 - (MI) SEQUENCE DESCRIPTION: SEQ ID NO:125:

Ala Gio Glo Ser Tie Ser Thr Xae Glo Xee lie Vol Pro 1 5

- (Z) INFORMATION FOR SEQ 10 NO:126:
 - (A) SHOURNCS CHARACTERISTICS: (A) LENGTH: 17 amino acids
 - (5) TYPE: smino sold
 - (C) STRANDERNESS:
 - (b) TOPOLOGY: Linear

143

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:

Amp Pro Glu Pro Mia Bro Pro Val Dro Thr Thr Ale Ale Ser Pro Pro I

Sec

- (2) INFORNATION FOR SEQ ID NO:127:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 15 amino acids
 - (B) TYPE: amino acid (C) STRANDEDWESS:
 - (D) TOPOLOGY: lisear
 - (21) ORQUENCE DESCRIPTION: SEQ 10 NO:127:

Als Pro Lys Thr Tyr Xaa Glu Giu Leo Lye Giy Thr Asp Thr Gly 1

- (2) INFORMATION FOR 980 10 NO.128:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 amino soids (B) TYPE: amino soid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: Minear
 - (x() SEGGENCE DESCRIPTION: SEG 15 No:188:

Asp Fit Ala Ser Ale Pro Asp Val Pro The Ale Ala Gle Los Thr Ser

Let ion Ash Ser Lot Ala Asp Pro Sen Val Sor Phe Ala Ash 25 30

- (2) INFORMATION FOR SEQ ID NO:129:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LANGTH: 22 amino acids
 - (B) TYPE: amino acid
 - (C) STRANGEDWESS:
 - (D) TOPOLOGI: Linear
 - (xx) SEQUENCE DESCRIPTION: SEQ ID NO:129:

Asp Pro Pre Asp Pro Hie Gin Was Asp Met Thr Lys Gly Tyr Tyr pro